

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 13:43:06 ; Search time 120 Seconds  
(without alignments)  
5799.237 Million cell updates/sec

Title: US-09-600-714-41  
Perfect score: 1254  
Sequence: 1 atgagcttaagtaccgcg.....attggctgttgatttaa 1254

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A-COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B-COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A-COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B-COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS-COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1185.2	94.5	1384	1	US-08-553-888A-1		Sequence 1, Appli
2	1182	94.3	1466	1	US-08-553-888A-2		Sequence 2, Appli
3	109.6	8.7	1650	4	US-09-461-325-75		Sequence 75, Appl
4	109.6	8.7	1650	4	US-10-012-542-75		Sequence 75, Appl
5	41	3.3	1884	4	US-09-621-976-14663		Sequence 14663, A
6	41	3.3	2094	4	US-09-252-991A-16118		Sequence 16118, A
7	41	3.3	2616	4	US-09-252-991A-16486		Sequence 16486, A
8	41	3.3	6909	4	US-09-199-637A-111		Sequence 111, App
9	40.6	3.2	1422	4	US-09-252-991A-8351		Sequence 8351, Ap
10	40.6	3.2	1503	4	US-09-252-991A-8488		Sequence 8488, Ap
11	40.6	3.2	1761	4	US-09-252-991A-8423		Sequence 8423, Ap
12	40.2	3.2	729	4	US-09-489-039A-5837		Sequence 5837, Ap
13	40	3.2	801	4	US-09-252-991A-6902		Sequence 6902, Ap
14	40	3.2	1410	4	US-09-252-991A-6871		Sequence 6871, Ap
15	40	3.2	1482	4	US-09-252-991A-6850		Sequence 6850, Ap
16	40	3.2	1500	4	US-09-252-991A-6757		Sequence 6757, Ap
17	39.4	3.1	467	2	US-08-476-176B-15		Sequence 15, Appl
18	39.4	3.1	467	3	US-08-127-721A-15		Sequence 15, Appl
19	39.4	3.1	467	3	US-08-485-246A-15		Sequence 15, Appl
20	39.4	3.1	468	2	US-08-476-176B-11		Sequence 11, Appl
21	39.4	3.1	468	3	US-08-127-721A-11		Sequence 11, Appl
22	39.4	3.1	468	3	US-08-485-246A-11		Sequence 11, Appl
23	39	3.1	1564	4	US-09-369-247-25		Sequence 25, Appl
24	38.8	3.1	4403765	3	US-09-103-840A-2		Sequence 2, Appli
25	38.8	3.1	4411529	3	US-09-103-840A-1		Sequence 1, Appli
26	38.2	3.0	2712	4	US-09-919-172-40		Sequence 40, Appl
27	38	3.0	1189	1	US-07-781-034-4		Sequence 4, Appli

28	38	3.0	1189	5	PCT-US92-08328-4	Sequence 4, Appli
29	37.8	3.0	500	4	US-09-370-838-142	Sequence 142, App
30	37.2	3.0	1384	4	US-09-372-422A-17	Sequence 17, Appl
31	36.6	2.9	897	4	US-09-668-680-5	Sequence 5, Appli
32	36.6	2.9	2031	4	US-09-475-515-60	Sequence 60, Appl
33	36.6	2.9	2553	4	US-09-475-515-66	Sequence 66, Appl
34	36.6	2.9	536165	4	US-09-214-808-1	Sequence 1, Appli
35	36.6	2.9	4403765	3	US-09-103-840A-2	Sequence 2, Appli
36	36.6	2.9	4411529	3	US-09-103-840A-1	Sequence 1, Appli
37	36.4	2.9	377	2	US-08-332-766A-1	Sequence 1, Appli
38	36.2	2.9	467	2	US-08-476-176B-17	Sequence 17, Appl
39	36.2	2.9	467	3	US-08-127-721A-17	Sequence 17, Appl
40	36.2	2.9	467	3	US-08-485-246A-17	Sequence 17, Appl
41	36.2	2.9	468	2	US-08-476-176B-13	Sequence 13, Appl
42	36.2	2.9	468	3	US-08-127-721A-13	Sequence 13, Appl
43	36.2	2.9	468	3	US-08-485-246A-13	Sequence 13, Appl
44	36.2	2.9	1941	4	US-09-489-039A-5585	Sequence 5585, Ap
45	36.2	2.9	2004	1	US-08-471-033-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1  
US-08-553-888A-1  
; Sequence 1, Application US/08553888A  
; Patent No. 5723293  
; GENERAL INFORMATION:  
; APPLICANT: Huang  
; TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/553.888A  
; FILING DATE: 11/06/95  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: O'Dea, Sean W.  
; REGISTRATION NUMBER: 37690  
; REFERENCE/DOCKET NUMBER: 454-5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1384 nucleotides  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-553-888A-1

Query Match	94.5%	Score 1185.2	DB 1	Length 1384
Best Local Similarity	96.6%	Pred. No. 0		
Matches 1211	Conservative	0	Mismatches 43	Indels 0
Gaps	0			
QY	1	ATGAGCTTAAGTACCGCGGTCTCTCGGGCTGCGCTGCGCCCTTGGCCCTTACACTG	60	
DB	31	ATGAGCTTAAGTACCGCGGTCTCTCGGGCTGCGCTGCGCCCTTGGCCCTTACACTG	90	
QY	61	GAAGCAGCTCTCATCTCTCTCTCTATTTTATCCCACTATGAGCTTCCCTTAGAGAT	120	

Db 91 GAAGCAGCTCTCATTTCTCTCTCTCTATTTTTTTTACCCACATAGAGCTTCTCTTAGAGAT 150  
Qy 121 CAAAAGGGGCTGTGGATCCTATCAAGTTGGCCAGAGATCTGACCGTGATGGCGGCATT 180  
Db 151 CAAAAGGGGCTGTGGATCCTATCAAGTCGGCCAGATCTGACCGTGATGGCGGCCTT 210  
Qy 181 GGGTTGGGCTTCTCACTCGAGTTTCCGGAGACACAGCTGGAGCGATGGGCTTCAAC 240  
Db 211 GGGTTGGGCTTCTCACTCAAAATTTCCGGAGACACAGCTGGAGCGATGGGCTTCAAC 270  
Qy 241 CTCTTCATGTGGCGCTTGGTGTGCGAGTGGGCAATCTGCTGGACGCTTCTCTGAGCCAG 300  
Db 271 CTCTTCATGTGGCGCTTGGTGTGCGAGTGGGCAATCTGCTGGACGCTTCTCTGAGCCAG 330  
Qy 301 TTCCCTTCTGGGAAGTGGTCATCACACTGTTGAGTATTCGGCTGGCCACCATGAGTGT 360  
Db 331 TTCCCTTCTGGGAAGTGGTCATCACACTGTTGAGTATTCGGCTGGCCACCATGAGTGT 390  
Qy 361 TTGTCGGTGTGATCTCAGTGGATGCTGCTTGGGGAAGTCAACTTGGGCGAGTTGGTG 420  
Db 391 ATGTCGGTGTGATCTCAGCGGGTGTGCTTCTGGGGAAGTCAACTTGGGCGAGTTGGTG 450  
Qy 421 GTGATGTGTGTGGAGTGACAGCTTTAGGCAACCTGAGGATGGTCAATGATATATC 480  
Db 451 GTGATGTGTGTGGAGTGACAGCTTTAGGCAACCTGAGGATGGTCAATGATATATC 510  
Qy 481 TTCAACACAGACTACACATGAATGATGATGATGATGATGATGATGATGATGATGATG 540  
Db 511 TTCAACACAGACTACACATGAATGATGATGATGATGATGATGATGATGATGATGATG 570  
Qy 541 CTGCTGTGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600  
Db 571 CTGCTGTGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 630  
Qy 601 ACAGCAAGATACCCAGTTGTGTGCAATGCTGGGCGCTCTTCTTGTGGATGTTTGG 660  
Db 631 AGAGCAAGATACCCAGTTGTGTGCAATGCTGGGCGCTCTTCTTGTGGATGTTTGG 690  
Qy 661 CCAAGTTTCACTCTGCTGCTGAGAGTCCATGCAAGGAGAGATGCGGTGTTCAAC 720  
Db 691 CCAAGTTTCACTCTGCTGCTGAGAGTCCATGCAAGGAGAGATGCGGTGTTCAAC 750  
Qy 721 ACCTACTATGCTGTAGCAGTCAGCGTGGTGACAGCCATCTCAGGGTCATCTTGGGTCC 780  
Db 751 ACCTACTATGCTGTAGCAGTCAGCGTGGTGACAGCCATCTCAGGGTCATCTTGGGTCC 810  
Qy 781 CCCCAAGGAGATACAGAGATTTATGTGCAAGTGGCGGTGTGGCAGAGCGGTGGCT 840  
Db 811 CCCCAAGGAGATACAGAGATTTATGTGCAAGTGGCGGTGTGGCAGAGCGGTGGCT 870  
Qy 841 GTGGGTACCTCGTGTACCTGATCCCTTCTCCGTGGCTTCCCATGGTGGTCTTGTG 900  
Db 871 GTGGGTACCTCGTGTACCTGATCCCTTCTCCGTGGCTTCCCATGGTGGTCTTGTG 930  
Qy 901 GCTGGGTGATCTCCGTGGGGGAGCAAGTACCTGCGGGGTGTTGTAAACCGAGTGTG 960  
Db 931 GCTGGGTGATCTCCGTGGGGGAGCAAGTACCTGCGGGGTGTTGTAAACCGAGTGTG 990  
Qy 961 GGGATTTCCACAGCTCCATCATGGGCTACAACTTCAAGCTTGGCTGGTGTGGAGAG 1020  
Db 991 GGGATTTCCACAGCTCCATCATGGGCTACAACTTCAAGCTTGGCTGGTGTGGAGAG 1050  
Qy 1021 ATCATCTACATTTGCTGTGCTGTGATACCGTGGAGCCGCAATGATGATGGC 1080  
Db 1051 ATCACCTACATTTGCTGTGCTGTGATCTGCTGAGACCGCAATGATGATGGC 1110  
Qy 1081 TTCCAGGTCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGTGATGATGATCTCAGCTCT 1140  
Db 1111 TTCCAGGTCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGTGATGATGATCTCAGCTCT 1170  
Qy 1141 GGTCTCTGACAGTTTCTCTCAATCTTAAATATGGAAGACCTCATGAGGCTAAA 1200  
Db 1171 GGTCTCTGACAGTTTCTCTCAATCTTAAATATGGAAGACCTCATGAGGCTAAA 1230

Qy 1201 TATTTTGATCACCAGTTTCTCGAAGTTTCTCTCAATTTGGCTGTGGATTTTAA 1254  
Db 1231 TATTTTGATCACCAGTTTCTCGAAGTTTCTCTCAATTTGGCTGTGGATTTTAA 1284

## RESULT 2

US-08-553-888A-2  
; Sequence 2, Application US/08553888A  
; Patent No. 5723293  
; GENERAL INFORMATION:  
; APPLICANT: Huang  
; TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR  
; DETERMINING RH BLOOD GROUP GENOTYPE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/553,888A  
; FILING DATE: 11/06/95  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: O'Dea, Sean W.  
; REGISTRATION NUMBER: 37690  
; REFERENCE/DOCKET NUMBER: 454-5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1466 nucleotides  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-553-888A-2

Query Match 94.3%; Score 1182; DB 1; Length 1466;  
Best Local Similarity 96.4%; Pred. No. 0;  
Matches 1209; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 1 ATGAGCTCTAAGTACCGGGTCTGTCGGCGCTGCTGCCCTCTGGGCGCTTAAACATG 60  
Db 45 ATGAGCTCTAAGTACCGGGTCTGTCGGCGCTGCTGCCCTCTGGGCGCTTAAACATG 104  
Qy 61 GAAGCAGCTCTCATTTCTCTCTATTTTACCACATATGACGCTTCTTTAGAGAT 120  
Db 105 GAAGCAGCTCTCATTTCTCTCTATTTTACCACATATGACGCTTCTTTAGAGAT 164  
Qy 121 CAAAAGGGGCTGTGGATCCTATCAAGTTGGCCAGAGATCTGACCGTGATGGCGGCATT 180  
Db 165 CAAAAGGGGCTGTGGATCCTATCAAGTTGGCCAGAGATCTGACCGTGATGGCGGCCTT 224  
Qy 181 GGGTTGGGCTTCTCACTCGAGTTTCCGGAGACACAGCTGGAGCGATGGGCTTCAAC 240  
Db 225 GGGTTGGGCTTCTCACTCAAAATTTCCGGAGACACAGCTGGAGCGATGGGCTTCAAC 284  
Qy 241 CTCTTCATGTGGCGCTTGGTGTGCGAGTGGGCAATCTGCTGCAACGGCTTCTTGAGCCAG 300  
Db 285 CTCTTCATGTGGCGCTTGGTGTGCGAGTGGGCAATCTGCTGCAACGGCTTCTTGAGCCAG 344  
Qy 301 TTCCCTTCTGGGAAGTGGTCATCACACTGTTGAGTATTCGGCTGGCCACCATGAGTGT 360  
Db 345 TTCCCTTCTGGGAAGTGGTCATCACACTGTTGAGTATTCGGCTGGCCACCATGAGTGT 404



RESULT 4  
 US-10-012-542-75  
 ; Sequence 75, Application US/10012542  
 ; Patent No. 6627741  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruben et al.  
 ; TITLE OF INVENTION: 94 Human Secreted Proteins  
 ; FILE REFERENCE: P2029P1  
 ; CURRENT APPLICATION NUMBER: US/10/012,542  
 ; CURRENT FILING DATE: 2001-12-12  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22  
 ; NUMBER OF SEQ ID NOS: 532  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO 75  
 ; LENGTH: 1650  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-012-542-75

Query Match 8.7%; Score 109.6; DB 4; Length 1650;  
 Best Local Similarity 48.7%; Pred. No. 4.7e-23;  
 Matches 327; Conservative 0; Mismatches 344; Indels 1; Gaps 1;  
 QY 364 TCGGTGCTGATCTCAGTGGATGCTCTTGGGAAGTCACTTGGCGCAGTTGGTGGTG 423  
 DB 33 TCTGCTGCTGGCTTTTGGGGAGTCTTGGTAAAGTCAGCCCATTCAGCTGCTCATC 92  
 QY 424 ATGTGCTGTGGAGTGACAGCTTTTAGGCAACCTGAGGATGTCATCAGTAATATCTTC 483  
 DB 93 ATGACTTTCTTCAAGTGACCTCTTCGTGTGAATGAGTTTCATCTCTTAACTGCTA 152  
 QY 484 AACACAGTACACATGACATGATGACATGACATGACATGACATGACATGACATGACAT 543  
 DB 153 AAGGTGAAGGATGACGAGGCTCCATGACCATCCACACATTTGGCGCTACTTTGGGCTC 212  
 QY 544 TCTGTGGCTGTGCTGCCAAAGCCCTTACCCGAGGAGGAGGATTAAGATCAGACA 603  
 DB 213 ACAGTGACCGGATCTCTACCGAGGACCACTAGACAGACAGACAGACAGACATCT 272  
 QY 604 GCAACGATACCAAGTTTGTGTCATGCTGGCGCCCTCTTTTGTGGATGTTCTGGCCA 663  
 DB 273 GTGTACCAAGTCGACCTTTTGGCCATGATTTGGCACCCTCTTCTGTGGATGTACTGGCC 332  
 QY 664 AGTTTCACTGCTGCTGAGAGTCCATCGAAGAGAGATGCGGTTCACACC 723  
 DB 333 AGCTTCAACTGAGCATATCCATCCATGGGACACCCAGCAGCCGCGCCATCAACACC 392  
 QY 724 TACTATGCTGTAGCAGTGTGACAGCCATCTCAGGTCATCTCTGGCTCACCCC 783  
 DB 393 TACTGCTCTTGGCAGCTGGTGTCTTACCTCGGTGGCAATATC-CAGTGCCTTCAAG 451  
 QY 784 CAAGGAGATGACAGACATGATGACAGTGGGTGTGGCAGGAGGCGTGGCTGTG 843  
 DB 452 AAGGCAAGCTGGAATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 511  
 QY 844 GGTACTGCTGTCACCTGATCTCTTCTCGGTGGTGTGTCATGTCATGTCATGTCATGTC 903  
 DB 512 GGTACCGTGTGAGATGATGCTCATGCTTACGGTGTCCCTCATCATCGGCTTGTGTGCT 571

QY 904 GGGCTGATCTCCGTGGGGAGGAGCAAGTACCTCCGGGGTGTGTATACCGAGTGTGGG 963  
 DB 572 GGCATCATCTCCACCCCTGGTTTGTATACCTGACCCCATTCCTGGAGTCCCGCTGCAC 631  
 QY 964 ATTCCCCACAGCTCCATCATGCGCTACAACTTCAGTTCGTGGTGTGCTTGGAGATC 1023  
 DB 632 ATCCAGGACACATGTGGCATTAAACAATCTGCATGGCATCTCTGGCATATAGCGGGCATC 691  
 QY 1024 ATCTACATGTG 1035  
 DB 692 GTGGGTGCTGTG 703  
 RESULT 5  
 US-09-621-976-14663  
 ; Sequence 14663, Application US/09621976  
 ; Patent No. 6639063  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; APPLICANT: Jobert, S.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
 ; FILE REFERENCE: GENSET.054PR2  
 ; CURRENT APPLICATION NUMBER: US/09/621,976  
 ; CURRENT FILING DATE: 2000-07-21  
 ; NUMBER OF SEQ ID NOS: 19335  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 14663  
 ; LENGTH: 484  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-621-976-14663

Query Match 3.3%; Score 41; DB 4; Length 484;  
 Best Local Similarity 74.6%; Pred. No. 0.018;  
 Matches 50; Conservative 1; Mismatches 16; Indels 0; Gaps 0;  
 QY 614 CCAGTTTGTCTGCATGCTGGGGCCCTCTTCTTGTGATGTTCTGCCCAAGTTTCACT 673  
 DB 416 CTTGTYGTCCACCACTAGGGGACCATCTTCTGTGATCTTCTGCTAGCTTCAATG 475  
 QY 674 CTGCTCT 680  
 DB 476 CTGCACT 482

RESULT 6  
 US-09-252-991A-16118/c  
 ; Sequence 16118, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 16118  
 ; LENGTH: 2094  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-16118

Query Match 3.3%; Score 41; DB 4; Length 2094;  
 Best Local Similarity 45.3%; Pred. No. 0.044;  
 Matches 149; Conservative 0; Mismatches 180; Indels 0; Gaps 0;  
 QY 770 CTTTGGCTCACCCCAAGGAGATCAGCAAGACTTATGTCACAGTCCGGTGTGGCAG 829

Db	1526	CCTCGGTACGCGCCCGCTGAGCTGGGCACCCCTGAGCTGCAGCGCCGAGATGTTGC	1767
Qy	830	GAGCGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGCTTGCATGCTGC	889
Db	1766	GCGTGCAGCCGAGGTGCGACCTTCTCAAGATTCCCCGCCGATCCGCCGTACCAAGG	1707
Qy	890	TGGGTCTGTGGCTGGGGTGATCTCCGTGGGGGAGCCAACTACTGCCGGGTGTTGTA	949
Db	1706	TGATTCGAGCOCCTGGCGACCAACAATCCTGTGTCGGCTGGCGTGCACCGGAC	1647
Qy	950	ACCGAGTCTGGGGATTCCCCACAGCTCCATCATGGGCTACAATTGAGCTTGGGTGC	1009
Db	1646	GCAGCAACCCGCCCCAAGCGTGCCTCGCCGCGGCGCGCTGCAGCGGGTCCGGCT	1587
Qy	1010	TGCTTGCAGAGATCATCTACAATTGTGTGTGGTGTGTGATACCGTCGGAGCGGCAATG	1069
Db	1586	CGCTGCGCCGGTTCATCCTGCTGTTGTGATGCTGGCGCAGACCTCGTGCACCTACT	1527
Qy	1070	GCATGATTGGCTTCAGGTCTCTCTCAGC	1098
Db	1526	ACATGAAGGCATCCTGCTCTACCAAGGC	1498

## RESULT 7

```

US-09-252-991A-16486
; Sequence 16486, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16486
; LENGTH: 2616
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16486

```

## RESULT 8

```

US-09-199-637A-111
; Sequence 111, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111
; LENGTH: 6909
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-111

```

## RESULT 9

;  
 US-09-252-991A-8351/c  
 ; Sequence 8351, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.1136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190



```
QY 357 TGGTTCGCTGCTGATCTCAGTGGATGCTGCTTTGGGAAGGTCAACTTGGCGCAGTT 416
Db 483 GGGTCTGCTGCTCAACTGCTGGCTGAAGAGCGGCGCTGATGTGGCGCGTGACCTATAT 542
QY 417 GGTGGTATGCTGCTGGTGGAGGTGACAGCTTTAGGCAACTGAGGATGTCATCAGTAA 476
Db 543 TGGGTGGTCTGCTTTGCTGGTCTGACGGGTATGACACCAAGAACTGAAAACATCGG 602
QY 477 TATCTTCAACACAGATACACATGACATGATGACATCTAGCTGTTGCGACGCTATTT 536
Db 603 CGAGCAGATTGATACCGCGACACGCGCAGCTGCTGCTAAATACTCGATCTCTGGCGCTT 662
QY 537 TGGGCTGCTGCTGGCCT 553
Db 563 AACGCTGATCTGGACT 679

RESULT 13
US-09-252-991A-6902/c
; Sequence 6902, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6902
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6902

Query Match 3.2%; Score 40; DB 4; Length 801;
Best Local Similarity 47.9%; Pred. No. 0.049;
Matches 115; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 236 TCAACCTCTTCATGCTGGCGCTTGGTGTGCAGTGGGCAATCTCTGTGGAGGTCACTTGGCGCAGT 295
Db 742 TCCTGCTCGGCCCTTGCACTTCCTGCTGGCGGATCAAGAGGCGCGCTTTCATCA 683
QY 296 GCAGTTCCTTCTGGGAAGGTGGTCATCACACTGTTCAAGTATTCGGCTGGCCACCATCA 355
Db 682 ACACCGTCACCAACCGTGGCCAGGTGTCGCGCTGTTCTGTTTCATCTGATCTGGCTGT 623
QY 356 GTGCTTTGCTGGTGTGATCTCAGTGGATGCTGTTTGGGGAAGGTCAACTTGGCGCAGT 415
Db 622 TCGCCTTCAAGCTGGACATCTTCACCGCGACATCTGGGCGAAGAGCAACCGGACCTGG 563
QY 416 TGGTGTGATGCTGCTGGTGGAGGTGACAGCTTTAGGCAACTGAGGATGTCATCAGTA 475
Db 562 GCAGCGTATGAACAGGTTGGCGCAACATGATGCTGCTACCGCTCTGGGTGTTTCATCGGCA 503

RESULT 14
US-09-252-991A-6871/c
; Sequence 6871, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6871
; LENGTH: 1480
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6871

Query Match 3.2%; Score 40; DB 4; Length 1482;
Best Local Similarity 47.9%; Pred. No. 0.072;
Matches 115; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 236 TCAACCTCTTCATGCTGGCGCTTGGTGTGCAGTGGGCAATCTCTGTGGAGGTCACTTGGCGCAGT 295
Db 417 TCCTGCTCTGGGCCCTTGCACTTCCTGCTGGCGGATCAAGAGGCGCGGTTCATCA 476
QY 296 GCAGTTCCTTCTGGGAAGGTGGTCATCACTGTTTCAAGTATTCGGCTGGCCACCATCA 355
Db 477 ACACCGTCACCAACCGTGGCCAAAGGTGCTGGCGCGCTTCTCTGTTTCATCTGATCTGGCTGT 536
QY 356 GTGCTTTGCTGGTGTGATCTCAGTGGATGCTGTTTGGGGAAGGTCAACTTGGCGCAGT 415
Db 537 TCGCCTTCAAGCTGACATCTTCACCGCGACATCTGGGCGAAGAGCAACCGGACCTGG 596
QY 416 TGGTGTGATGCTGCTGGTGGAGGTGACAGCTTTAGGCAACTGAGGATGTCATCAGTA 475
Db 597 GCAGCGTATGAACAGGTTGGCGCAACATGATGCTGGTTCACCGTCTGGGTGTTTCATCGGCA 656
```

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6871
; LENGTH: 1410
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6871

Query Match 3.2%; Score 40; DB 4; Length 1410;
Best Local Similarity 47.9%; Pred. No. 0.07;
Matches 115; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 236 TCAACCTCTTCATGCTGGCGCTTGGTGTGCAGTGGGCAATCTCTGTGGAGGTCACTTGGCGCAGT 295
Db 1046 TCCTGCTCTGGGCCCTTGCACTTCCTGCTGGCGGATCAAGAGGCGCGGTTCATCA 987
QY 296 GCAGTTCCTTCTGGGAAGGTGGTCATCACACTGTTTCAAGTATTCGGCTGGCCACCATCA 355
Db 986 ACACCGTCACCAACCGTGGCCAAAGGTGCTGGCGCGCTTCTCTGTTTCATCTGATCTGGCTGT 927
QY 356 GTGCTTTGCTGGTGTGATCTCAGTGGATGCTCTTGGGGAAGGTCACTTGGCGCAGT 415
Db 926 TCGCCTTCAAGCTGGACATCTTCACCGCGACATCTGGGCGAAGAGCAACCGGACCTGG 867
QY 416 TGGTGTGATGCTGCTGGTGGAGGTGACAGCTTTAGGCAACTGAGGATGTCATCAGTA 475
Db 866 GCAGCGTATGAACAGGTTGGCGCAACATGATGCTGTGTACCGCTCTGGGTGTTTCATCGGCA 807

RESULT 15
US-09-252-991A-6850
; Sequence 6850, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6850
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6850

Query Match 3.2%; Score 40; DB 4; Length 1482;
Best Local Similarity 47.9%; Pred. No. 0.072;
Matches 115; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 236 TCAACCTCTTCATGCTGGCGCTTGGTGTGCAGTGGGCAATCTCTGTGGAGGTCACTTGGCGCAGT 295
Db 417 TCCTGCTCTGGGCCCTTGCACTTCCTGCTGGCGGATCAAGAGGCGCGGTTCATCA 476
QY 296 GCAGTTCCTTCTGGGAAGGTGGTCATCACTGTTTCAAGTATTCGGCTGGCCACCATCA 355
Db 477 ACACCGTCACCAACCGTGGCCAAAGGTGCTGGCGCGCTTCTCTGTTTCATCTGATCTGGCTGT 536
QY 356 GTGCTTTGCTGGTGTGATCTCAGTGGATGCTGTTTGGGGAAGGTCAACTTGGCGCAGT 415
Db 537 TCGCCTTCAAGCTGACATCTTCACCGCGACATCTGGGCGAAGAGCAACCGGACCTGG 596
QY 416 TGGTGTGATGCTGCTGGTGGAGGTGACAGCTTTAGGCAACTGAGGATGTCATCAGTA 475
Db 597 GCAGCGTATGAACAGGTTGGCGCAACATGATGCTGGTTCACCGTCTGGGTGTTTCATCGGCA 656
```

Search completed: August 25, 2004, 17:30:45  
Job time : 130 secs





Publication No. US20030073623A1	
GENERAL INFORMATION:	
APPLICANT: Hyseq, Inc.	
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED	
FROM VARIOUS CDNA LIBRARIES	
FILE REFERENCE: 20411-756	
CURRENT APPLICATION NUMBER: US/09/918,995	
CURRENT FILING DATE: 2001-07-30	
PRIOR APPLICATION NUMBER: US/09/235,076	
PRIOR FILING DATE: 1999-01-20	
NUMBER OF SEQ ID NOS: 38054	
SOFTWARE: FastSeq for Windows Version 3.0	
SEQ ID NO 30694	
LENGTH: 487	
TYPE: DNA	
ORGANISM: Homo sapiens	
FEATURE:	
NAME/KEY: misc.feature	
LOCATION: (1)...(487)	
OTHER INFORMATION: n = A,T,C or G	
US-09-918-995-30694	
Query Match 29.2%; Score 365.8; DB 10; Length 487;	
Best Local Similarity 99.5%; Pred. No. 5.3e-107;	
Matches 367; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	706 AATCCGCTGTTCAACACCTACTATGTGTAGCAGTCAGCGTGTGTGACAGCCATCTCAGGG 765
Db	56 ATTCCCGTGTTCACACCTACTACTATGTGTAGCAGTCAGCGTGTGTGACAGCCATCTCAGGG 115
QY	766 TCATCTCTGGCTCACCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGGGTGTG 825
Db	116 TCATCTCTGGCTCACCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGGGTGTG 175
QY	826 GCAGGAGCGTGTGTGTGGTACTCTGTGTCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 885
Db	176 GCAGGAGCGTGTGTGTGGTACTCTGTGTCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 235
QY	886 GTCTGGTCTTGTGGCTGGCTGATCTCGTGGGGGAGCCCAAGTACCTGCCGGGTGT 945
Db	236 GTCTGGTCTTGTGGCTGGCTGATCTCGTGGGGGAGCCCAAGTACCTGCCGGGTGT 295
QY	946 TGTAAACCGAGTGTGGGATTCGCCACAGCTCCATCATGGGCTACAACTTCAGCTTCTG 1005
Db	296 TGTAAACCGAGTGTGGGATTCGCCACAGCTCCATCATGGGCTACAACTTCAGCTTCTG 355
QY	1006 GGTCTGCTGGAGAGATCATCTACATTTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 1065
Db	356 GGTCTGCTGGAGAGATCATCTACATTTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 415
QY	1066 AATGGCATG 1074
Db	416 AATGGCATG 424
RESULT 3	
US-10-302-172-362	
Sequence 362, Application US/10302172	
Publication No. US20040053250A1	
GENERAL INFORMATION:	
APPLICANT: Tang, Y. Tom	
APPLICANT: Xue, Aidong J.	
APPLICANT: Drmanac, Radoje T.	
TITLE OF INVENTION: NO. US20040053250A1el Arginine-rich Protein-like Nucleic Acids	
TITLE OF INVENTION: Polypeptides	
FILE REFERENCE: 803.1CNCP	
CURRENT APPLICATION NUMBER: US/10/302,172	
CURRENT FILING DATE: 2002-11-21	
PRIOR APPLICATION NUMBER: US 10/225,251	
PRIOR FILING DATE: 2002-08-20	
PRIOR APPLICATION NUMBER: PCT US02/05095	
PRIOR FILING DATE: 2002-03-05	
PRIOR APPLICATION NUMBER: US 09/799,451	

241	CTCTTCATCTGCGCTGTGTGTGAGTGGGCAATCTCTGTGACGCTTCTTGACCCAG 300
Db	241 CTCTTCATCTGCGCTGTGTGTGAGTGGGCAATCTCTGTGACGCTTCTTGACCCAG 300
QY	301 TTCCCTCTGGGAGTGTGTATCACACTGTTTCAGTATTTCGGTGGCCACCATGAGTGTCT 360
Db	301 TTCCCTCTGGGAGTGTGTATCACACTGTTTCAGTATTTCGGTGGCCACCATGAGTGTCT 360
QY	361 TTCTCGGTGTGTATCTCAGTGGATGTCTTCTGGGAAGTCAACTTGGCGCATTTGGTG 420
Db	361 TTCTCGGTGTGTATCTCAGTGGATGTCTTCTGGGAAGTCAACTTGGCGCATTTGGTG 420
QY	421 GTGATGGTGTGTGTGAGGTGACAGCTTTAGGCAACTGAGGATGTGTATCATCAATATC 480
Db	421 GTGATGGTGTGTGTGAGGTGACAGCTTTAGGCAACTGAGGATGTGTATCATCAATATC 480
QY	481 TTCAACACAGTACACATGACATGATGACATCTACGTGTTCGAGCTATTTTGGG 540
Db	481 TTCAACACAGTACACATGACATGATGACATCTACGTGTTCGAGCTATTTTGGG 540
QY	541 CTGTCTGTGGCTGTGTGCTCCAAAGCTCTTACCCGAGGGAACGGAGGATAAAGATCAG 600
Db	541 CTGTCTGTGGCTGTGTGCTCCAAAGCTCTTACCCGAGGGAACGGAGGATAAAGATCAG 600
QY	601 ACAGCAACGATACCCAGTTTGTCTGCCATCTGTGGGCGCCCTCTTCTGTGGATGTTCTGG 660
Db	601 ACAGCAACGATACCCAGTTTGTCTGCCATCTGTGGGCGCCCTCTTCTGTGGATGTTCTGG 660
QY	661 CCAAGTTTCAACTCTGTCTGTGAGAACTTCAATCGAAAGGAAGATGCGGTGTTCAAC 720
Db	661 CCAAGTTTCAACTCTGTCTGTGAGAACTTCAATCGAAAGGAAGATGCGGTGTTCAAC 720
QY	721 ACCTACTATCTGTAGCAGTGTGTGTGAGCAAGCTATCTCAGGATCATCTCTGGCTCAC 780
Db	721 ACCTACTATCTGTAGCAGTGTGTGTGAGCAAGCTATCTCAGGATCATCTCTGGCTCAC 780
QY	781 CCCCAAGGAGATCAGCAAGACTTATGTGACAGTGGGTGTGTGGCAGAGGGCTGGCT 840
Db	781 CCCCAAGGAGATCAGCAAGACTTATGTGACAGTGGGTGTGTGGCAGAGGGCTGGCT 840
QY	841 GTGGGTACCTGTGTACCTGTATCCCTTCTCCGTGGCTTGCATGTGTGTGTGTGTGTGTGT 900
Db	841 GTGGGTACCTGTGTACCTGTATCCCTTCTCCGTGGCTTGCATGTGTGTGTGTGTGTGTGT 900
QY	901 GCTGGGTGTGTCTCCGTGGGAGCAAGTACTGTCCGGGTGTGTAAACCGAGTGTGT 960
Db	901 GCTGGGTGTGTCTCCGTGGGAGCAAGTACTGTCCGGGTGTGTAAACCGAGTGTGT 960
QY	961 GGGATTTCCACAGCTCCATCATGGGTACAACTTCACTTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Db	961 GGGATTTCCACAGCTCCATCATGGGTACAACTTCACTTGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY	1021 ATCATCTACATTTGTGTGTGTGTGTGTATACCGTTCGAGCCGCAATGGCATGTTGGC 1080
Db	1021 ATCATCTACATTTGTGTGTGTGTGTATACCGTTCGAGCCGCAATGGCATGTTGGC 1080
QY	1081 TTCAGGTCTTCCTCAGATTTGGGAACTCAGTGTGGCCATCTGATGCTCTCAGCTCT 1140
Db	1081 TTCAGGTCTTCCTCAGATTTGGGAACTCAGTGTGGCCATCTGATGCTCTCAGCTCT 1140
QY	1141 GGTCTCCTGACAGTGTGTCTCTAAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200
Db	1141 GGTCTCCTGACAGTGTGTCTCTAAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200
QY	1201 TATTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGTGATTTTAA 1254
Db	1201 TATTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGTGATTTTAA 1254
RESULT 2	
US-09-918-995-30694	
Sequence 30694, Application US/09918995	

; PRIOR FILING DATE: 2001-03-05  
 ; NUMBER OF SEQ ID NOS: 950  
 ; SOFTWARE: pt FL\_genes Version 2.0  
 ; SEQ ID NO 362  
 ; LENGTH: 1792  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (32)..(1354)  
 US-10-302-172-362

Query Match 14.0%; Score 175.4; DB 13; Length 1792;  
 Best Local Similarity 49.4%; Pred. No. 3.2e-45;  
 Matches 455; Conservative 0; Mismatches 466; Indels 0; Gaps 0;

QY	141	CTATCAAGTTGGCCAAAGATCTGACCGTATGCGCGCCATTGGCTGGCTTCCTCACTC 200
DB	211	CTACCCAGCTTCCAGGACGTGATGCCATGGTCTTGCGGGCTTTGACTTCTCATGGT 270
QY	201	GAGTTTCGGAGACACAGCTGGAGCAGTGTGGCCCTTCAACCTCTTTCATGCTGGCGCTGG 260
DB	271	CTTCCTGCAGCGTTACGGCTTCAGCAGCGTGGCTTCACCTTCTCTGGCGCTTTC 330
QY	261	TGTCAGTGGCCATCTCTGAGCGCTTCTGAGCGAGTTCCTCTTCTGGAGAGTGGT 320
DB	331	CCTCAGTGTGCCACTGTGTCCAGGGCTTCTCCACTCTCTCCACGGTGGCCACATCCA 390
QY	321	CATCACACTGTTCAAGTATTCGGCTGGCCACCATGAGTGTCTTGGTGGTGTGATCTCAGT 380
DB	391	TGTTGGCGTGGAGAGCATGATCAATGTGACTTTTGTGGGGGCGGTCTCATCTCCTT 450
QY	381	GGATGCTGTCTTGGGGAAGGTCACTTTGGCGAGTGGTGGTGTGATGGTGGTGGAGGT 440
DB	451	TGTTGGCGTGGAGAGCATGATCAATGTGACTTTTGTGGGGGCGGTCTCATCTCCTT 510
QY	441	GACAGCTTTAGGACACCTGAGATGGTATCATGATTAATCTTCAACACAGCTACCAAT 500
DB	511	GGTGTCTTTGGCATCAATGAGTTTGTGCTCTTTCATCTCTCTGGGGGTGAGATGCGGG 570
QY	501	GAACATCATGACATCTACGCTGTTCGACGCTATTTTGGGCTGTCTGTGGCTTGGTCT 560
DB	571	AGGTCATGATATCCACACTTTGGTGGCTTCTCGGGCTGGTCTCTTTCGGGGGTCT 630
QY	561	GCCAAAGCTCTACCCGAGGGAAGAGGAGTAAAGATCAGACAGCAACGATACCCAGTTT 620
DB	631	GTACAGGCCCGAGCTGAGAGAGCAAGCAGCCGCGGCTCGTCTACCATTCAGACCT 690
QY	621	GTCTGCCATGCTGGGCGCCCTCTCTTGTGATGTTCTGGCAAGTTTCAACTCTGTCT 680
DB	691	CTTCGCCATGATTTGGGACCACTCTCTCTGTGATCTTCTGGCTAGCTTCAATGTGCACT 750
QY	681	GCTGAGAAGTCCAATCAAGAGGAAGTACCGGTGTTTCAACACCTACTATCTCTAGCAGT 740
DB	751	CAGAGCGCTGGGGCTGGGAGCATCGAGCGGCCCTCAACATATCTCTCCCTGGCTGC 810
QY	741	CAGCGTGGTGAAGCCATCTCAGGCTCATCTTGGCTCACCCCAAGGGGAAGATCAGCAA 800
DB	811	CAGCACCCCTTGGCACCTTTTCCCTTGTGAGCCCTGTAGGGGAAGATGGAGGCTTGACAT 870
QY	801	GACTTATGTGACAGTCCGCTGTGAGAGGGGTGGTGTGGTACCTCTGTCGCTCACT 860
DB	871	GGTCCACATCAAAATGACGCTGGCTGGGAGGGGTGTGGTGGGACCTTCAAGTGAAT 930
QY	861	GATCCCTTCTCCGCTGGCTTGCATGCTGGTGTGCTGTGGCTGGCTGATCTCCGCTCG 920
DB	931	GATGCTGACACCTTTGGGGCTCTGGAGCTGGCTTCTTGGCTGGGACTGTCTCCACGCT 990
QY	921	GGGAGCCAGTACTGCTCCGGGGTGTGTAACCGAGTGTCTGGGATCCCCACAGCTCCAT 980
DB	991	GGGGTACAAGTCTTTCACGCCCATCTTGAATCAAAATTCAAAGTCCAGACACATGTGG 1050
QY	981	CATGGGGTCAACTTCAGCTTGTGGGTGTGCTTGGAGATCATCTACATTTGTGCTGT 1040

Db 1051 AGTCCACAACTCCATGGGATGCGGGGTCCTCTGGGGGCTCTCTCTGGGGTCTCTGGG 111  
 QY 1041 GGTGCTTGATACCGTCGGAGC 1061  
 Db 1111 TGGACTTGGCACCACCATGAAGC 1131

RESULT 4  
 US-10-074-978A-41/C  
 ; Sequence 41, Application US/10074978A  
 ; Publication No. US20040010119A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leite, Mario  
 ; APPLICANT: Spytek, Kimberly A  
 ; APPLICANT: Guo, Xiaojia (Sasha)  
 ; APPLICANT: Fernandes, Elma  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Liu, Xiahong  
 ; APPLICANT: Casman, Stacie  
 ; APPLICANT: Boldog, Ferenc  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Bialock, Angela  
 ; APPLICANT: Ballinger, Robert  
 ; APPLICANT: Vernet, Corine  
 ; APPLICANT: Tchernev, Velizar T  
 ; APPLICANT: Malyankar, Uriel M  
 ; APPLICANT: Gusev, Vladimir  
 ; APPLICANT: Rastelli, Luca  
 ; APPLICANT: Mezes, Peter S  
 ; APPLICANT: Ellerman, Karen  
 ; APPLICANT: Heyes, Melvin P  
 ; APPLICANT: Herrman, John  
 ; APPLICANT: Pena, Carol E A  
 ; APPLICANT: Shimkets, Richard A  
 ; APPLICANT: Taupier Jr, Raymond J  
 ; APPLICANT: Moore, No. US20040010119A11le  
 ; APPLICANT: Sheno, Suresh  
 ; APPLICANT: Edinger, Shlomit  
 ; APPLICANT: Gunther, Erik  
 ; APPLICANT: Stone, Dave  
 ; APPLICANT: Millet, Isabelle  
 ; APPLICANT: Peyman, John  
 ; APPLICANT: Smithson, Glennda  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: 21402-269  
 ; CURRENT APPLICATION NUMBER: US/10/074,978A  
 ; CURRENT FILING DATE: 2003-01-07  
 ; PRIOR APPLICATION NUMBER: 60/268,221  
 ; PRIOR FILING DATE: 2001-02-12  
 ; PRIOR APPLICATION NUMBER: 60/335,109  
 ; PRIOR FILING DATE: 2001-10-31  
 ; PRIOR APPLICATION NUMBER: 60/312,284  
 ; PRIOR FILING DATE: 2001-08-14  
 ; PRIOR APPLICATION NUMBER: 60/268,496  
 ; PRIOR FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/276,703  
 ; PRIOR FILING DATE: 2001-03-16  
 ; PRIOR APPLICATION NUMBER: 60/330,293  
 ; PRIOR FILING DATE: 2001-10-18  
 ; PRIOR APPLICATION NUMBER: 60/322,127  
 ; PRIOR FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/280,899  
 ; PRIOR FILING DATE: 2001-04-02  
 ; PRIOR APPLICATION NUMBER: 60/310,797  
 ; PRIOR FILING DATE: 2001-08-08  
 ; PRIOR APPLICATION NUMBER: 60/268,646  
 ; PRIOR FILING DATE: 2001-02-14  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 547  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 41

```
; LENGTH: 1351
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-074-978A-41

Query Match      13.2%; Score 165.4; DB 16; Length 1351;
Best Local Similarity 49.1%; Pred. No. 4.6e-42; Indels 0; Gaps 0;
Matches 439; Conservative 0; Mismatches 456;

QY 141 CTATCAAGTTGGCCAAAGATCTGACCGTGATGGCGCCATTGGCTTGGGCTTCTCCACCTC 200
Db 1175 CTACCCAAAGCTTCCAGGACGTGCACGTGATGTCTTCTGGTGGCTTCCGCTTCTCATGAC 1116

QY 201 GAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAACCTTTCATGCTGGCGCTTGG 260
Db 1115 TTTTCTCGACGCTGACGGCTTCAAGCCGCTGGGCTTCAACTTCTGTGGCAGCTTCCG 1056

QY 261 TGTGCAAGTGGCAATCTCTGCTGACAGGCTTCTTGAGCCAGTTCCTTCTGGGAAGTGGT 320
Db 1055 CATCCAGTGGCGCTGCTCATGACAGGCTGGTTCCACTTCTTACAAGACCGCTACATCGT 996

QY 321 CATCACACTGTTCAAGTATTCGGTGGCCACCATGAGTGTCTTGTGGTGTGATCTCAGT 380
Db 995 CGTGGCGGTGGAGAACCTCATCAACGCTGACTTCTGGTGGCTCTGTCTGCGTGGCCTT 936

QY 381 GAGTGTCTGTTGGGAGAGTCAACTTGGCGCAGTTGGTGGTGTGATGGTGGTGGAGT 440
Db 935 TGGGGCAGTGTCTGGGTAAAGTCAAGCCCATTCAGCTGCTCATCATGACTTCTTCCAACT 876

QY 441 GACAGCTTTTAGGCAACCTCGAGGATGTGTCATCAGTAATATCTTCAACACAGACTACCACAT 500
Db 875 GACCTCTTCGCTGTGAATGAGTTTCACTTCTTAACTGCTAAAGTGAAGATGCAGG 816

QY 501 GAAATGATGACATCTAGTGTTCGAGCCATATTTTGGGCTGTCTGTGGCTTGGTGGCT 560
Db 815 AGGCTCCATGACCATCACACATTTGGCGCTACTTTGGGCTCACAGTACCCGGATCCT 756

QY 561 GCCAAAGCTCTACCCGAGGAGCGAGGATGAAGATCAGACAGCAACGATACCCAGTTT 620
Db 755 CTACCGACGCAACCTAGAGAGAGCAAGAGGAGAGAGAGAAATCTGTGACAGTGGACCT 696

QY 621 GTCTGCCATGTGGCGCCCTCTTCTTGTGGATGTTCTGGCCAAAGTTTCAACTCTGCTCT 680
Db 695 CTTTGGCATGATTGGCACCTCTTCTTGTGGATGTAAGTCTGGCCAGCTTCAACTCAGCCAT 636

QY 681 GCTGAGAAGTCCAAATCGAAAGGAAGATGCGGTGTTCAACACCTACTATGCTGTAGCAGT 740
Db 635 ATCCTACCATGGGAGACAGCCAGCACCGAGCGGCCATCAACACCTTACTGCTCTTGGCAGC 576

QY 741 CAGGTGTGACAGCCATCTCAGGCTCATCTTGGCTCACCCCAAGGAGATCAGCAA 800
Db 575 CTGGTGTCTTACCTCGGTGGCAATATCAAGTGGCTCTGACAGAGGCGAGCTGGACAT 516

QY 801 GACTTATGTGACAGTGGTGTGGAGGAGCGGTGGCTGTGGGTACCTGCTGTACCT 860
Db 515 GGTGCACATCCAGAAATCCAGCTCGCAGGAGGGGTGGCGGTGGGTACCGTGTGAGAT 456

QY 861 GATCCCTTCTCGTGGCTTGCATGTTGCTGGGTCTTGTGGCTGGGTGATCTCCGTGG 920
Db 455 GATGCTCATGCTTTACGTGGCTCTCATCATCGGTCTCTGTGGCGCATCATCTCCACCT 396

QY 921 GGGAGCCAAAGTACCTGCGGGGTGTGTAAACCGAGTGTGGGATTCGCCACAGCTCCAT 980
Db 395 GGGTTTGTATACCTGACCCCATCTCGGAGTCCGGCTGCACATCAGGACACATGTGG 336

QY 981 CATGGCTTACAACTTACGCTGTGGGTCTGCTTGGAGAGATCATCTACATTTGG 1035
Db 335 CATTAACAATCTGCAATGCAATTCCTGGCATCATAGGGGCGCATCGTGGGTGCTGTG 281
```

RESULT 5  
US-09-949-145-5  
; Sequence 5, Application US/09949145

```
; Patent No. US20020055622A1
; GENERAL INFORMATION:
; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyco
; FILE REFERENCE: Docket 454-31
; CURRENT APPLICATION NUMBER: US/09/949,145
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1440
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-949-145-5

Query Match      13.2%; Score 165.4; DB 9; Length 1440;
Best Local Similarity 49.1%; Pred. No. 4.8e-42; Indels 0; Gaps 0;
Matches 439; Conservative 0; Mismatches 456;

QY 141 CTATCAAGTTGGCCAAAGATCTGACCGTGATGGCGCCATTGGCTTGGGCTTCTCCACCTC 200
Db 177 CTACCCAAAGCTTCCAGGACGTGCACGTGATGTCTTCTGGGCTTCCGCTTCTCATGAC 236

QY 201 GAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAACCTTCTCATGTGGCGCTTGG 260
Db 237 TTTCTTGCAGCGTACGGCTTCAAGCCGCTGGGCTTCAACTTCTGTGGCAGCTTCCG 296

QY 261 TGTGCAAGTGGCAATCTCTGCTGGACCGCTTCTTGAGCCAGTTCCTTCTGGGAAGTGGT 320
Db 297 CATCCAGTGGCGCTGCTCATGACGGCTTGGTTCCTTCTTACAAGACCGCTACATCGT 356

QY 321 CATCACACTGTTCAAGTATTCGGCTGGCCACCATGAGTGTCTTCTGGCTGTGATCTCAGT 380
Db 357 CGTGGCGTGGAGAACTCATCAACGCTGACTTCTGGTGGCTCTGTCTGGCTGGCTT 416

QY 381 GAGTGTCTCTTGGGAAAGGTCAACTTGGCGCAGTGTGGTGGTGTGATGCTGTGTGGAGT 440
Db 417 TGGGGCAGTCTCTGGTAAAGTCAAGCCCATTCAGCTCATCATGACTTCTTCTTCCAAGT 476

QY 441 GACAGCTTTAGGCAACCTGAGGATGGTCACTAGTATATCTTCAACACAGACTACCACAT 500
Db 477 GACCTCTTTCGCTGTGAATGAGTTCATCTTCTTAACTCTTAAAGTGAAGATGACGAG 536

QY 501 GAAATGATGACATCTACGTTTTCGAGCCTATTTTGGCTGTCTTGGCTGTCTGGCTGCT 560
Db 537 AGGCTCATGACCATCCACATTTGGCGCTTACTTTGGGCTCACAGTGCACCGGATCCT 596

QY 561 GCCAAAGCTCTTACCCGAGGAAACGGAGGATAAAGATCAGACAGCAACGATACCCAGTTT 620
Db 597 CTACCGACGCAACCTAGAGCAGAGCAAGGAGAGACAGAAATCTGTGTACAGTGGGACCT 656

QY 621 GTCTGCCATGTGGGCGCCCTCTTCTTGTGATGTTCTGGCCAAAGTTTCAACTCTGCTCT 680
Db 657 CTTTGGCATGATTGGCACCTCTTCTTGTGATGTACTTGGCCAGCTTCAACTCAGCCAT 716

QY 681 GCTGAGAAGTCCAAATCGAAAGGAGAAATGCGGTCTTCAACACCTACTATGCTGTAGCAGT 740
Db 717 ATCTACCATGGGAGACAGCCAGCCAGCGGCCATCAACCTTCTGCTTCTTGGCAGC 776

QY 741 CAGCGTGTGACAGCCATCTCAGGCTCATCTTGGCTCATCCCCCAAGGAGATCAGCAA 800
Db 777 CTGCGTCTTACCTCGGTGGCAATATCCAGTGGCTCTGACAAAGAGGCAAGCTGGACAT 836

QY 801 GACTTATGTGACAGTGTGGTGTGGAGGAGCGGTGGCTGTGGTACCTGCTGCTACCT 860
Db 837 GGTGCACATCCAGAAATCCAGCTCGCAGGAGGGGTGGCGGTGGGTACCGTGTGAGAT 896

QY 861 GATCCCTTCTCCGTGGCTTGCATGTTGCTGGGTCTTGTGGCTGGGTGATCTCCGTGG 920
Db 897 GATGCTCATGCTTACGGTGGCTTCTGCTGGCTTCTGTGGCGCATCATCTCCACCT 956
```

QY 921 GGGAGCCAAAGTACCTGCGGGGTCTTTGTAACCGAGTCTCGGGATTCGCCACAGCTCCAT 980  
Db |||||  
QY 957 GGGTTTGTATCTGACCCATCTCTGGAGTCCCGCTGCAATCCAGACATGTGG 1016  
Db |||||  
QY 981 CATGGGCTACAATCTCAGCTTGTGGTCTCTGCTTGGAGAGATCATCTPACATTTGTG 1035  
Db |||||  
1017 CATTAAATCTGATGGCATCTCTGGCATCATAGGGGCATCGTGGGTCTGTG 1071  
RESULT 6  
US-09-949-145-1  
; Sequence 1, Application US/09949145  
; Patent No. US20020055622A1  
; GENERAL INFORMATION:  
; APPLICANT: New York Blood Center  
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc  
; FILE REFERENCE: Docket 454-31  
; CURRENT APPLICATION NUMBER: US/09/949,145  
; PRIORITY FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: US 60/230660  
; PRIOR FILING DATE: 2000-09-07  
; NUMBER OF SEQ ID NOS: 77  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1952  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: AF193809  
; DATABASE ENTRY DATE: 1999-12-22  
; RELEVANT RESIDUES: (1)..(1952)  
US-09-949-145-1

Query Match 13.2%; Score 165.4; DB 9; Length 1952;  
Best Local Similarity 49.1%; Pred. No. 5.7e-42;  
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;  
QY 141 CTATCAAGTTGGCCAAAGATCTGACCGTGTGAGTGGGCTTGGGCTTCTCTCACCTC 200  
Db |||||  
QY 201 CTATCAAGTTGGCCAAAGATCTGACCGTGTGAGTGGGCTTGGGCTTCTCTCACCTC 260  
Db |||||  
QY 201 GAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAACCTCTTCATGCTGGGCTTGG 260  
Db |||||  
QY 261 TGTGAGTGGGCAATCTCTGCTGGAGCGGCTTCCCTGAGCAGTTCCTTCTGGGAAGGTGGT 320  
Db |||||  
QY 321 CATCAGTGGGCGCTGCTCATGAGGGCTGGTTCCACTTCTTACAGACCGCTACATCGT 380  
QY 321 CATCAGTGGGCGCTGCTCATGAGGGCTGGTTCCACTTCTTACAGACCGCTACATCGT 380  
Db |||||  
QY 381 CGTGGCGTGGAGAACCTCATCAACGCTGACTTCTGGGTGGCTCTGTCTGGTGGCTT 440  
QY 381 GGAATCTGCTTGGGAGAGTCAATGTCGCGAGTGTGGTGTGATGCTGTGGTGGAGT 440  
Db |||||  
QY 441 TGGGCGAGTCTGGGTAAGTCAAGCCCATTCAGCTGCTCATGACTTCTTCCAAAT 500  
QY 441 GACAGCTTTAGCAACCTGAGATGCTATCAGTAATATCTTCAACAGACAGTACCAT 500  
Db |||||  
QY 501 GACCTCTTCGTGTGAATGAGTTCACTTCTCTTAACTGCTTAAAGGTGAAGGATGACAG 560  
QY 501 GAAATATGACATCTAGCTTTCGACGCTTATTTGGGCTGTCTGTGGCTGGTGGCT 560  
Db |||||  
QY 561 AGGCTCCATGACCATCCACATTTGGGCTTCTTGGGCTTCAAGTACCCGATCTT 620  
QY 561 GCCAAGCTCTAACCCGAGGAGACGAGATTAAGATCAGACAGAACATACCCAGTTT 620  
Db |||||  
QY 621 CTACCGACGCAACCTTAGAGCAGAGACAGAGACAGAAATCTGTGTACCACTGGACCT 680  
QY 621 GTCTCCATGCTGGGCGCTCTCTCTGTGGATGTTCTGGCCAACTTCAACTCTGTCT 680  
Db |||||  
QY 681 CTTTGCCATGATGGCAACCTCTCTCTGTGGATGACTGGGCCAGCTTCAACTCAGCAAT 740

QY 681 GCTGAGAGTCCAAATCGAAAGGAAGAAATGCGGTGTTCAACACCTACTATCTCTAGCAGT 740  
Db |||||  
QY 741 ATCTTACCATGGGAGACAGCCAGCAGCGAGCGGCATCAACACCTACTGCTCTTGGCAGC 800  
Db |||||  
QY 741 CAGCGTGGTGACAGCCATCTCAGGGTCATCTTGGCTCACCCCAAGGGAAGATCAGCAA 800  
Db |||||  
QY 801 CTGCGTGTCTTACTCTCGTGGCAATATCCAGTGCCTTGCAACAAGGGCAAGGTGGACAT 860  
QY 801 GACTTATGTGCAACAGTCCGCTGTGGCAGGAGCGTGGCTGTGGGTACCTCTGCTCACCT 860  
Db |||||  
QY 861 GGTGCATCCAGATGCCAGCTCGCAGGAGGGTGGCGTGGGTACCGCTGCTGAGAT 920  
QY 861 GATCCCTTCTCGTGGCTTCCCATGTGCTGGTCTTGTGGCTGGGCTGATCTCCGTCGG 920  
Db |||||  
QY 921 GATGCTCATGCTCTTACGGTGGCTCATCATCGGCTTCTGTGGGCTCATCTCCACCT 980  
QY 921 GGGAGCAAAGTACTCGCCGGGTGTGTAACCGAGTGTGGGATTCGCCACAGCTCCAT 980  
Db |||||  
QY 981 GGGTTTGTATACCTGACCCCATTCCTGGAGTCCCGGCTGCATCCAGACACATGTGG 1040  
Db |||||  
QY 1041 CATTAAATCTGCATGGCATCTCTGGCATCATAGCGGCACTGCTGGTCTGTG 1095

## RESULT 7

US-10-342-887-1635  
; Sequence 1635, Application US/10342887  
; Publication No. US20040058340A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter S.  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Christopher J.  
; APPLICANT: Van 't Veer, Laura Johanna  
; APPLICANT: Van de Vijver, Marc J.  
; APPLICANT: Bernards, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-188-999  
; CURRENT APPLICATION NUMBER: US/10/342,887  
; PRIOR FILING DATE: 2003-01-15  
; PRIOR APPLICATION NUMBER: 60/298,918  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/380,710  
; PRIOR FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: 10/172,118  
; PRIOR FILING DATE: 2002-06-14  
; NUMBER OF SEQ ID NOS: 2699  
; SEQ ID NO 1635  
; LENGTH: 1952  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-342-887-1635

Query Match 13.2%; Score 165.4; DB 13; Length 1952;  
Best Local Similarity 49.1%; Pred. No. 5.7e-42;  
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCAAAGATCTGACCGTGTGAGTGGGCTTGGGCTTCTCTCACCTC 200  
Db |||||  
QY 201 CTATCAAGTTGGCCAAAGATCTGACCGTGTGAGTGGGCTTGGGCTTCTCTCACCTC 260  
Db |||||  
QY 201 GAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAACCTCTTCATGCTGGGCTTGG 260  
Db |||||  
QY 261 TGTGAGTGGGCAATCTCTGCTGGAGCGGCTTCCCTGAGCAGTTCCTTCTGGGAAGGTGGT 320  
Db |||||  
QY 321 CATCAGTGGGCGCTGCTCATGAGGGCTGGTTCCACTTCTTACAGACCGCTACATCGT 380  
QY 321 CATCAGTGGGCGCTGCTCATGAGGGCTGGTTCCACTTCTTACAGACCGCTACATCGT 380  
Db |||||

```

381 CGTGGCGTGGGAACCTCATCAACGCTGACTTCTCGTGGCTCTGTCTGGTGGCTT 440
381 GGATGCTGTCTTGGGGAAGTCAACTTTGGCGAGTTGGTGGTGGTGGTGGTGGTGGT 440
441 TGGGGCAGTCTTGGGTAAGTCAAGCCATTCAGCTGCTCATCATGACTTCTTCCAAAT 500
441 GACAGCTTTAGGCAACCTGAGGATGCTCATGATATATCTTCAACACAGACTTACCACAT 500
501 GACCTCTTGGCTGTAAGTGGTCAATCTCTTAACTGCTTAAAGTGAAGATGAGG 560
501 GAACATGATGACATCTACGCTGTTGCGAGCTTATTTTGGCTCTCTGTGGCTTGGTGCCT 560
561 AGCTCCATGACCATCCACATTTGGCGCTTACTTTGGCTCACAGTGAACCGGATCTCT 620
561 GCCAAGCTCTTACCGAGGGAACGGAGGTAAGATCAGACAGCAAGTATCCAGTTT 620
621 CTACCGAGCAACCTAGAGCAGAGCAAGGAGACAGAAATCTGTGTACAGTCCGACCT 680
621 GTCTGCCATGCTGGCGCTCTTCTTGTGGATGTTCTGGCCAAAGTTTCAACTCTGCTCT 680
681 CTTTGGCATGATTGGCAACCTCTTCTGTGTAAGTACTGGCCAGCTTCAACTCAGCCAT 740
681 GCTGAGAAGTCCAAATCGAAAGGAAGATGCCGTGTTCAACACCTACTATGCTGTAGCAGT 740
741 ATCTACCATGGGAGACAGCCAGCAGCGCCATCAACACCTACTGCTCTTGGCAGC 800
741 CAGGTGGTGAAGCCATCTCAGGCTCATCTTGGCTCAGCCGAGGGAAGATCAGCAA 800
801 CTGCGTCTTACCTCGTGGCAATATCCAGTCCCTGCACAAGAGGGAAGCTGGACAT 860
801 GACTTATGTCACAGTCCGCTTGGCAGAGGCTGGCTGTGGTACCTCTGTACCT 860
861 GGTGCAATCCAGATCCAGCTCGCAGAGGCTGGCTGTGGTACCTGCTGTAGAT 920
861 GATCCCTTCTCCGTGGCTGCGATGCTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGG 920
921 GATGCTCATGCTTACGCTGCTCTCATCGGCTCTGTGGTCTGTGGTCTGTGGTCTGTGG 980
921 GGGAGCCAGTACTGCTGGGCTGTGTAAACGAGTCTGGGATTCACACAGCTCCAT 980
981 GGGTTTGTATACCTGACCTCCCATTCCTGGAGTCCCGCTGCACATCAGGACATGTGG 1040
981 CATGGGTACAACTTCAGCTTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGG 1035
1041 CATTAACAATCTGCATGCAATCTCTGGCATCTAGGCGGATCTGGTGTGTG 1095

```

```

RESULT 8
US-10-172-118-1635
; Sequence 1635, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380, 770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1635
; LENGTH: 1952
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_016321

```

```

; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1635
Query Match 13.2%; Score 165.4; DB 13; Length 1952;
Best Local Similarity 49.1%; Pred. No. 5.7e-42; Indels 0; Gaps 0;
Matches 439; Conservative 0; Mismatches 456;
141 CTATCAAGTTGGCCAAAGATCTGACCGTGTATGGCGGCATTTGGCTTGGGCTTCTCACTC 200
201 CTATCCAAAGCTTCCAGACGTCGACGTCGATGCTTCTGTTGGGCTTCCGCTTCTCATGAC 260
201 GAGTTTCCGAGACACAGCTGAGAGAGTGTGGCTTCAACCTCTTCACTGCTGGCTTGG 260
261 TTTCTCTGACGCGTACGGCTTACGCGCTTGGGCTTCAACTTCTTGGCAGCTTCCG 320
261 TGTGCACTGGGCAATCTCTGCTGAGCGCTTCTTGAGCCAGTTCCTTCTTGGGAAGTGGT 320
321 CATCCAGTGGGCGCTCTCATGAGGCTGTGTTCCACTTCTTACAAGACCGCTACATCGT 380
321 CATCACACTGTTTCAATTCGGCTGGCCACCACTGATGCTTGTTCGGTGTGATCTCACT 380
381 CGTGGCGTGGGAACCTCATCAACGCTGACTTCTGCTGGCTCTGTCTGCTGGCTT 440
381 GGATGCTGCTTGGGGAAGTCAACTTGGCGAGTGTGGTGGTGGTGGTGGTGGTGGTGGT 440
441 TGGGCGAGTCTGGGTAAAGTCAAGCCATTCAGCTGCTCATCATGACTTCTTCCAGT 500
441 GACAGCTTTAGGCAACCTGAGGATGTCATCAGTAATATCTTCAACACAGACTTACCACAT 500
501 GACCTCTTGGCTGTAAGTGGTCAATCTCTTAACTGCTTAAAGTGAAGATGAGG 560
501 GAACATGATGACATCTACGCTTTCGAGCTTATTTTGGCTGTGTGGCTTGGTGGTGGT 560
561 AGCTCCATGACCATCCACATTTGGCGCTTACTTTGGCTCACAGTACCGGATCTCT 620
561 GCCAAGCTCTTACCGAGGGAACGGAGGTAAGATCAGACAGCAAGTATCCAGT 620
621 CTACCGAGCAACCTAGAGCAGAGCAAGGAGACAGAAATCTGTGTACAGTCCGACCT 680
621 GTCTGCCATCTGGGCGCTCTTCTTGTGGATGTTCTGGCCAAAGTTTCAACTCTGCTCT 680
681 CTTTGCCATGATTTGGCAACCTCTTCTGTGGATGTTCTGGCCAGCTTCAACTCAGCCAT 740
681 GCTGAGAAGTCCAAATCGAAAGGAAGATGCGCTTCAACACCTACTATGCTGTAGCAGT 740
741 ATCTACCATGGGAGACAGCCAGCAGCCAGCGCCATCAACACCTACTGCTCTTGGCAGC 800
741 CAGGTGGTGAAGCCATCTCAGGCTCATCTTGGCTCAGCCCAAGGGAAGATCAGCAA 800
801 CTGCTGCTTACCTCGTGGCAATATCCAGTCCCTGCACAAGAGGCAAGCTGGACAT 860
801 GACTTATGTCACAGTCCGCTTGGCAGAGGCTGGCTGTGGTACCTCTGTCTCACT 860
861 GGTGCAATCCAGATCCAGCTCGCAGAGGCTGGCTGTGGTACCTGCTGTAGAT 920
861 GATCCCTTCTCCGTGGCTTGCATGCTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGG 920
921 GATGCTCATGCTTACGCTGCTCTCATCGGCTCTCATCGGCTCTGTCTGCGGCATCATCTCCACCT 980
921 GGGAGCCAGTACTGCTGGGCTGTGTAAACGAGTCTGGGATTCACACAGCTCCAT 980
981 GGGTTTGTATACCTGACCTCCCATTCCTGGAGTCCCGCTGCACATCAGGACATGTGG 1040
981 CATGGGTACAACTTCACTGCTTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGG 1035
1041 CATTAACAATCTGCATGCAATCTCTGGCATCTAGGCGGATCTGGTGTGTG 1095

```

```

RESULT 9
US-09-949-145-6
; Sequence 6, Application US/09949145
; Patent No. US2002005562A1
; GENERAL INFORMATION:

```

```

; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc
; FILE REFERENCE: Pocket 454-31
; CURRENT APPLICATION NUMBER: US/09/949,145
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-949-145-6

Query Match      12.0%; Score 150.2; DB 9; Length 1497;
Best Local Similarity 47.5%; Pred. No. 4e-37;
Matches 446; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

QY  97  CACTATGAGCTTCCTTAGAGGATCAAAAGGGGCTCGTGGCATCTTCCACCTAGGATTTGCATACCTA 156
Db  136  CGCAAGAAATCTCCAGGACGTTGAGACGAGTTTACTATGCGTACCGGAGCTTCCAG 195
QY  157  GATCTGACCGTGATGGCGGCATTTGGCTTGGCTTCTCACTCGAGTTTCCGGAGACAC 216
Db  196  GATGTACACGCCATGTCTTGTGGGCTTCGGCTTCTCATGACCTCTCTGAGCGCTAC 255
QY  217  AGCTGAGCAGTGTGGCTTCAACCTCTTCACTGCTGGCGCTTGGGTGAGTGGGCAATC 276
Db  256  GCGCTTCAGCGCTGTAGGCTTCACTTCTGCTGGCAGCTTTCGGCATCCAGTGGGCACTG 315
QY  277  CTGCTGGAGCGCTTCTCGAGCAGTTCCTCTCTGGAAGGTGGTCACTACATGTTCACT 336
Db  316  CTCATGACGAGTGTTCATTACTTTGAAGAGGCCCATTTGTCCTGAGCGTCGAGAAC 375
QY  337  ATTGCGCTG3CCACCAATGATGTTTGTGCGTGCTGATCTCAGTGGATGCTGTCTTGGG 396
Db  376  ATCATCCAAAGTGACTTCTGTGTGGCATCTTCTGTGTGGCTTCGGGCGAGTTCTAGGC 435
QY  397  AAGGTCACTTGGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 456
Db  436  AAGGTCAAGCCGATGCGCTGCTCATTAATGACCTTCTTCCAACTGACTCTCTTACAGTG 495
QY  457  CTGAGGATGCTCATCAGTAATATCTTCAACAGACTACCAATCAATGATGACATC 516
Db  496  ATGAGTTCATCTCTGAACTGTATAGGCAAGAGATGAGGGGCTCTATGACCAATC 555
QY  517  TACGTTTGGCAGCTATTTTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 576
Db  556  CACACATTTGGGCGCTACTTTGGGCTCAAGTGACCTGGATCTCTACCGAATAACCTG 615
QY  577  GAGGAAACGGAGATAAAGATCAGACAGCAAGATACCCAGTTTGTCTGCCATGTGGGC 636
Db  616  GATCAGAGCAACAGACAGAGCTCAGTGTATACCACTCGGACCTTTTGGCCATGATGGC 675
QY  637  GCGCTCTTCTGTGATGTTCTGGGCAAGTTTCAACTCTGCTCTGCTGAGAGTCCAATC 696
Db  676  ACCCTCTTCTGTGATATATGCGCCAGTTTCAATTCAGCCAGTTTCTTCCACGGAGAT 735
QY  697  GAAAGGAAGATGCGGTGTTCAACACCTACTATGCTGTAGCAGTCAAGGCTGTGACAGCC 756
Db  736  GCGGACCGCAGCAGCGCTCAATACCTACTCTCTCTTGGCAGCGAGTGTCTAACCA 795
QY  757  ATCTCAGGGTCACTCTTGGCTCACCCCAAGGAAGATCAGCAAGACTTATGTGACAGT 816
Db  796  GTGACAGTATCAGTAATTGTACACAAGAAAGGCAAGTTGGATATGCTGACATCCAGAT 855
QY  817  GCGGTGTTGGCAGGAGCGTGTCTGTGGTACTCTGTGCTACCTGCTACCTGATCCCTTCTCCGTGG 876
Db  856  GCCAGCTTGCAGTGGGGTGGGTGTGGGCAAGCTGCGGAGATGATGCTCACACCTTAC 915
QY  877  CTTGCCATGCTGGGTCTTGTGGCTGTGGCTGTGATCTCCGTTCGGGGGAGCCAAAGTACCTG 936

```

```

Db  916  GGGGCTCTCATCGTGGGTTCTTCTCGGCAATTTTCTCCACCTAGGATTTGCATACCTA 975
QY  937  CCGGGGTGTTGTAACCGAGTGTCTGGGATTTCCCCACAGATCTCCATCATGGCTCAAACTTC 996
Db  976  ACGCCATTCTCGAGTCCCGCTTCGCATCCAGGACACATGGGCAATTCACAACTGCAC 1035
QY  997  AGCTTCTCGGTCTGCTTGGAGAGATCATCTACATTTGTG 1035
Db  1036  GGCATTCCTGGCATCATAGCGGCATTTGGGTGCTGTG 1074

RESULT 10
US-09-949-145-2
; Sequence 2, Application US/09949145
; Patent No. US20020055622A1
; GENERAL INFORMATION:
; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc
; FILE REFERENCE: Pocket 454-31
; CURRENT APPLICATION NUMBER: US/09/949,145
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2097
; TYPE: DNA
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AF193810
; DATABASE ENTRY DATE: 1999-12-22
; RELEVANT RESIDUES: (1)..(2097)
US-09-949-145-2

Query Match      12.0%; Score 150.2; DB 9; Length 2097;
Best Local Similarity 47.5%; Pred. No. 4.8e-37;
Matches 446; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

QY  97  CACTATGAGCTTCCTTAGAGGATCAAAAGGGGCTCGTGGCATCTTCACTCAAGTTGGCCAA 156
Db  258  CGCAAGAAATCTCCAGGACGTTGAGAACGAGTTTACTATCGCTACCGAGCTTCAG 317
QY  157  GATCTGACCGTATGGCGGCCATTTGGCTTGGGCTTCTTCACTCGAGTTTCCGGAGACAC 216
Db  318  GATGTACACGCCATGCTTCTGTGGGCTTCGGCTTCTCTCATGACCTTCTCTGAGCGCTAC 377
QY  217  AGCTGAGCAGTGTGGCTTCAACCTCTTCACTGCTGGCGCTTGGTGTGAGTGGGCAATC 276
Db  378  GCGTTCAAGCTGTAGGCTTCAACTTCTGCTGGCAGCTTTGGGATCCAGTGGGCACTG 437
QY  277  CTGCTGAGCGGCTTCTGAGCGAGTTCCCTTCTGGGAAGGTGGTCACTACACTGTTCACT 336
Db  438  CTCATGCGAGGATGTTTCCATTACTTTGAAGAGCCACATTTGCTGAGCGTCAGAAC 497
QY  337  ATTGCGTGGGCAACATGAGTGTCTTGTGCGTGTGATCTCAGTGGATGCTGTCTTGGGG 396
Db  498  ATCATCAAGCTGACTTCTGTGTGGCATCTTCTGTGTGGCTTCTGCGGCGAGTTCTTAGGC 557
QY  397  AAGGTCAACTTGGCGAGTTGGTGTGATGCTGTGTGGGAGTGAACAGCTTTAGGCAAC 456
Db  558  AAGGTCAAGCGGATGAGCTGCTCATATGACCTTCTTCCAAAGTACTCTCTTCAAGTGTG 617
QY  457  CTGAGGATGCTCAGTAAATATCTTCAACACAGACTACCAATGAACATGATGACATC 516
Db  618  AATGAGTTCATCTCTCCATCACTCATAGAGCAAGAGATGAGGGGCTCTATGACCAATC 677
QY  517  TACGTGTTGGCAGCTATTTTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 576
Db  678  CACACATTTGGCGCTACTTTGGGCTCACAGTGTGAGTGTGATCTCTACCGAAAAACCTG 737
QY  577  GAGGGAACGGAGGATAAGATCAGACAGCAAGATACCCAGTTTGTCTGCCATGCTGGGC 636

```

Db 738 GATCAGCAAGACAGACAGACAGCTCAGTGTACCACTCGGACCTTTTGGCCATGATTGGC 797  
QY 637 GCCCTCTCTTGTGGAGTGTTCGGCAAGTTTCAACTCTGCTCTGCTGTGAGAGTCAATC 696  
Db 798 ACCCTCTCTTGTGGATATACCTGCCAGTTTCAATTCAGCCAGTTTCCCTTCCACGGAGAT 857  
QY 697 GAAAGGAAGATGCCGTGTTCAACACCTACTATGCTGTAGCAGTCAAGCGTGTGACAGCC 756  
Db 858 GCCCAGACCGAGACGCCCTCAATACCTACTCTCTCTTGGCAGCGAGTGTGCTAACACA 917  
QY 757 ATCTCAGGGTCACTCTTGGCTACCCCCCAAGGAAGATCAGCAAGACTTATGTGCACAGT 816  
Db 918 GTGACAGTATCCAGTATGTACACAAGAGGCAAGTTGGATATGTGTGCACATCCAGAAT 977  
QY 817 GCGGTGTGACAGAGCGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 876  
Db 978 GCCACGTTTGCAGTGGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1037  
QY 877 CTGTCATGCTGTGGTCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 936  
Db 1038 GCGCTCTCATCTGTGGGTTCTTCTGCGGCAATTTCTCCACCTAGGATTTGCATACCTA 1097  
QY 937 CCGGGGTGTTTAAACCGAGTGTGGGATTTCCCAAGCTTCCATCATGCGGCTCAACTTC 996  
Db 1098 ACGCCATTCTCGAGTCCCGCTTGCATCCAGGACACATGTGCGATTCAACACCTGCAC 1157  
QY 997 AGCTTGTGGTCTGCTGTGAGAGATCATCTACATGTG 1035  
Db 1158 GGCATTCTGCGCATCATAGCGCGCATTTGTGGTGTGTG 1196

RESULT 11  
US-10-104-047-730/c  
; Sequence 730, Application US/10104047  
; Publication No. US20030236392A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20030236392A1 full length cDNA  
; FILE REFERENCE: HI-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; PRIOR FILING DATE: 2002-03-25  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 730  
; LENGTH: 2415  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-104-047-730

Query Match 11.9%; Score 148.8; DB 16; Length 2415;  
Best Local Similarity 65.9%; Pred. No. 1.5e-36;  
Matches 216; Conservative 0; Mismatches 112; Indels 0; Gaps 0;  
QY 758 TCTCAGGGTCACTCTGCTCACCCTCCCAAGGAAGATCAGCAAGACTTATGTGCACAGT 817  
Db 1875 TCTCTCTACCTTGTCTTCTTACCAACAGCTATTTCTTTCGACACTTATGTGCACAGT 1816  
QY 818 CGGTGTTGGCAGGAGCGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 877  
Db 1815 CGGTGTTGGCAGGAGCGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1756  
QY 878 TTGCCATGCTGGTCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 937  
Db 1755 TTGCCATGCTGGTCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1696  
QY 938 CGGGTGTGTAAACCGAGTGTGGGATTTCCCAAGCTTCCATCATGCGGCTCAACTTCA 997  
Db 1695 CGGTAAAGAACTAGACAACCTACCTCTCTGCTTTGGCTGAAGGCCAGCAGACGCTGGG 1636  
QY 998 GCTTGTCTGGTCTGCTGTGAGAGATCATCTACATGTGCTGTGCTGTGCTGTGCTGTGCT 1057

Db 1635 ACCTGATGGCCCACTGTGCAGTGCACAGCTAGGAGGTGTGCGGCGCATTTCTTTA 1576  
QY 1058 GAGCGGCAATGSCATGATTGCTTCCA 1085  
Db 1575 TTGGCTTCAACGCCTAGTAGGAGATCCA 1548  
RESULT 12  
US-10-027-632-282039  
; Sequence 282039, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: Polymorphisms in the Human Genome  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 282039  
; LENGTH: 505  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-282039

Query Match 11.5%; Score 144.6; DB 13; Length 505;  
Best Local Similarity 86.9%; Pred. No. 1.4e-35;  
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
QY 758 TCTCAGGGTCACTCTTGGCTCACCCTCCCAAGGAAGATCAGCAAGACTTATGTGCACAGT 817  
Db 241 TCTCTCTACCTTGTCTTCTTACCCACACGCTATTTCTTTCGACACTTATGTGCACAGT 300  
QY 818 CGGTGTTGGCAGGAGCGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 877  
Db 301 CGGTGTTGGCAGGAGCGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 360  
QY 878 TTGCCATGCTGGGTCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 937  
Db 361 TTGCCATGCTGGGTCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 420  
QY 938 CGG 940  
Db 421 CGG 423

RESULT 13  
US-10-027-632-282040  
; Sequence 282040, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: Polymorphisms in the Human Genome  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006



```

; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282040
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282040

Query Match      11.5%; Score 144.6; DB 13; Length 505;
Best Local Similarity 86.9%; Pred. No. 1.4e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 758 TCTCAGGGTCATCCTTGGCTCACCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTG 817
Db 241 TCTCTCTACCTTGCCTTCTTTACCCACAGCTATTTCTTTGGCAGACTTATGTGCACAGTG 300

QY 818 CGGTGTTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 877
Db 301 CGGTGTTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 360

QY 878 TTGCCATGTCGTGGGTCTTGTGGCTGGGTGATCTCCGTCGGGGGAGCAAGTACTCTGC 937
Db 361 TTGCCATGTCGTGGGTCTTGTGGCTGGGTGATCTCCGTCGGGGGAGCAAGTACTCTGC 420

QY 938 CGG 940
Db 421 CGG 423

RESULT 14
US-10-027-632-282039
; Sequence 282039, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282039
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282040

Query Match      11.5%; Score 144.6; DB 13; Length 505;
Best Local Similarity 86.9%; Pred. No. 1.4e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 758 TCTCAGGGTCATCCTTGGCTCACCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTG 817
Db 241 TCTCTCTACCTTGCCTTCTTTACCCACAGCTATTTCTTTGGCAGACTTATGTGCACAGTG 300

QY 818 CGGTGTTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 877
Db 301 CGGTGTTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 360

QY 878 TTGCCATGTCGTGGGTCTTGTGGCTGGGTGATCTCCGTCGGGGGAGCAAGTACTCTGC 937
Db 361 TTGCCATGTCGTGGGTCTTGTGGCTGGGTGATCTCCGTCGGGGGAGCAAGTACTCTGC 420

QY 938 CGG 940
Db 421 CGG 423

RESULT 15
US-10-027-632-282040
; Sequence 282040, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282040
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282040

Query Match      11.5%; Score 144.6; DB 16; Length 505;
Best Local Similarity 86.9%; Pred. No. 1.4e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 758 TCTCAGGGTCATCCTTGGCTCACCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTG 817
Db 241 TCTCTCTACCTTGCCTTCTTTACCCACAGCTATTTCTTTGGCAGACTTATGTGCACAGTG 300

QY 818 CGGTGTTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 877
Db 301 CGGTGTTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 360

QY 878 TTGCCATGTCGTGGGTCTTGTGGCTGGGTGATCTCCGTCGGGGGAGCAAGTACTCTGC 937
Db 361 TTGCCATGTCGTGGGTCTTGTGGCTGGGTGATCTCCGTCGGGGGAGCAAGTACTCTGC 420

QY 938 CGG 940
Db 421 CGG 423
```

```

US-10-027-632-282039

Query Match      11.5%; Score 144.6; DB 16; Length 505;
Best Local Similarity 86.9%; Pred. No. 1.4e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 758 TCTCAGGGTCATCCTTGGCTCACCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTG 817
Db 241 TCTCTCTACCTTGCCTTCTTTACCCACAGCTATTTCTTTGGCAGACTTATGTGCACAGTG 300

QY 818 CGGTGTTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 877
Db 301 CGGTGTTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 360

QY 878 TTGCCATGTCGTGGGTCTTGTGGCTGGGTGATCTCCGTCGGGGGAGCAAGTACTCTGC 937
Db 361 TTGCCATGTCGTGGGTCTTGTGGCTGGGTGATCTCCGTCGGGGGAGCAAGTACTCTGC 420

QY 938 CGG 940
Db 421 CGG 423

RESULT 15
US-10-027-632-282040
; Sequence 282040, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282040
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282040

Query Match      11.5%; Score 144.6; DB 16; Length 505;
Best Local Similarity 86.9%; Pred. No. 1.4e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 758 TCTCAGGGTCATCCTTGGCTCACCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTG 817
Db 241 TCTCTCTACCTTGCCTTCTTTACCCACAGCTATTTCTTTGGCAGACTTATGTGCACAGTG 300

QY 818 CGGTGTTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 877
Db 301 CGGTGTTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 360

QY 878 TTGCCATGTCGTGGGTCTTGTGGCTGGGTGATCTCCGTCGGGGGAGCAAGTACTCTGC 937
Db 361 TTGCCATGTCGTGGGTCTTGTGGCTGGGTGATCTCCGTCGGGGGAGCAAGTACTCTGC 420

QY 938 CGG 940
Db 421 CGG 423
```

Db           |||  
              421 CGG 423

Search completed: August 25, 2004, 18:55:15  
Job time : 625 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 05:21:23 ; Search time 522 seconds

(without alignments)

10205.441 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254

Sequence: 1 atgagctctaagtaaccgcg.....atttgctgttgattttaa 1254

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001s:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002s:\*
- 7: Geneseqn2003s:\*
- 8: Geneseqn2003bs:\*
- 9: Geneseqn2003cs:\*
- 10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1254	100.0	1254	2	AAX86522
2	1252.4	99.9	2811	9	ADE09785
3	365.8	29.2	487	8	ACH43482
4	177	14.1	1805	6	ACH43482
5	175.4	14.0	1792	6	ABX12333
6	165.4	13.2	1351	7	ACD20392
7	165.4	13.2	1853	3	AAC64286
8	165.4	13.2	1949	5	AH25763
9	165.4	13.2	1952	6	ABK49215
10	165.4	13.2	1968	7	ABX76257
11	165.4	13.2	1970	3	AAC76344
12	159.4	12.7	1679	6	ABK12334
13	152.4	12.2	3706	4	AH18728
14	150.2	12.0	2098	6	ABK49216
15	148.8	11.9	2415	9	ADB62576
16	147.8	11.8	399	9	ADB62576
17	139.8	11.1	486	4	AAI11615
18	139.8	11.1	486	4	ABA53307
19	139.8	11.1	486	4	AAI32911
20	139.8	11.1	486	4	ABA42886
21	139.8	11.1	486	4	ABA23081
22	139.8	11.1	486	4	AAK27012
23	139.8	11.1	486	4	AAK01568

C	24	139.8	11.1	486	4	ABS26596	Abs26596 Human liv
C	25	139.8	11.1	486	5	AAI01542	AAI01542 Probe #15
C	26	139.8	11.1	486	6	ABS01595	Abs01595 Human gen
C	27	135	10.8	1188	9	ADE07444	Ade07444 Novel cod
C	28	120	9.6	1571	6	ABL90445	Abl90445 Human pol
C	29	113.8	9.1	1840	4	ABL02133	Abl02133 Drosophil
C	30	111.8	8.9	123	4	AAI20829	AAI20829 Probe #10
C	31	111.8	8.9	123	4	ABA65900	Aba65900 Human foe
C	32	111.8	8.9	123	4	AAI46064	AAI46064 Probe #14
C	33	111.8	8.9	123	4	ABA48015	Aba48015 Human bre
C	34	111.8	8.9	123	4	ABA32986	Aba32986 Probe #11
C	35	111.8	8.9	123	4	AAK40045	Aak40045 Human bon
C	36	111.8	8.9	123	4	AAK14315	Aak14315 Human bra
C	37	111.8	8.9	123	4	ABS39627	Abs39627 Human liv
C	38	111.8	8.9	123	5	AAI06540	AAI06540 Probe #65
C	39	111.8	8.9	123	6	ABS14127	Abs14127 Human gen
C	40	109.6	8.7	1650	3	AAZ97083	Aaz97083 Human sec
C	41	109.6	8.7	1650	7	ADA56347	Ada56347 Gene enco
C	42	109.6	8.7	1650	7	ADA40187	Ada40187 Human sec
C	43	109.6	8.7	1650	8	ACH66712	Ach66712 Novel hum
C	44	109.6	8.7	1650	9	ADC73726	Adc73726 Human sec
C	45	109.6	8.7	1650	9	ADD37696	Add37696 Human sec

## ALIGNMENTS

RESULT 1  
AAX86522  
ID AAX86522 standard; cDNA; 1254 BP.  
AC AAX86522;  
XX  
XX  
DT 04-OCT-1999 (first entry)  
DE cDNA sequence of the prevalent allele of the Rhd gene.  
XX  
KW Allele; Rhesus D antigen; Rhd; weak D phenotype; blood transfusion; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1254  
FT /\*tag= a  
XX  
PN WO9937763-A2.  
XX  
PD 29-JUL-1999.  
XX  
PF 18-DEC-1998; 98WO-EP008319.  
XX  
PR 23-JAN-1998; 98EP-00101203.  
XX  
PA (DRKB-) DRK BLUTSPENDEDIENST BADEN WUERTEMBERG.  
XX  
PI Flegel WA, Wagner FF;  
XX  
DR WPI; 1999-469127/39.  
XX  
DR P-PSDB; AAY24056.  
XX  
PT Nucleic acid sequences correlated with Rhesus weak D phenotype, useful for screening blood from donors and recipients for transfusion methods.  
XX  
PS Disclosure; Fig 2; 64pp; English.  
XX  
CC The present sequence represents the prevalent allele of the Rhesus D (Rhd) antigen gene. The specification describes a Rhd contributing to or indicative of the weak D phenotype, where the Rhd polynucleotide carries at least one missense mutation as compared to the wild-type Rhd, in its transmembrane and/or intracellular regions, especially in amino acid positions 2-16, 114-149, 179-225 or/and 267-397, with the proviso that the D antigen does not carry a single missense mutation leading to a F223V or T283I substitution. The probes and antibodies are useful in the

CC methods for detection of weak D phenotypes. Red blood cells, from  
CC probands, are useful for the assessment of the affinity, avidity and/or  
CC reactivity of monoclonal anti-D antibodies, polyclonal anti-D antisera or  
CC of anti-globulin or anti-human-globulin antisera. Detecting the presence  
CC of the RhD associated with weak D phenotype is useful for determining  
CC that a patient in need of a blood transfusion is to be transfused with  
CC RhD negative blood from a donor. Alternatively, testing for weak D  
CC phenotype RHD in the blood of a donor is useful for determining whether  
CC the donor blood should be excluded for transfusion to patients having  
CC wild type RHD or weak D types, other than that of the donor weak D type  
XX  
SQ Sequence 1254 BP; 241 A; 322 C; 351 G; 340 T; 0 U; 0 Other;

Query Match 100.0%; Score 1254; DB 2; Length 1254;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCTCTAAGTACCGCGGTCTGTCGGCGCTGCTGCCCTCTGGGCGCTTAAACACTG 60  
DB 1 ATGAGCTCTAAGTACCGCGGTCTGTCGGCGGTGCTGCCCTCTGGGCGCTTAAACACTG 60  
QY 61 GAAGCAGCTCTCAATCTCTCTCTCTATTTTAAACCACTATGACGCTTCTTTAGAGGAT 120  
DB 61 GAAGCAGCTCTCAATCTCTCTCTCTATTTTAAACCACTATGACGCTTCTTTAGAGGAT 120  
QY 121 CAAAGGGGCTCGTGGATCTATCAAGTTGGCCAGATCTGACCGTGTGGCGGCATT 180  
DB 121 CAAAGGGGCTCGTGGATCTATCAAGTTGGCCAGATCTGACCGTGTGGCGGCATT 180  
QY 181 GGCTTGGGCTCTCACTCGAGTTTCCGGAGACACAGCTGGAGCAGTGGCCCTTCAAC 240  
DB 181 GGCTTGGGCTCTCACTCGAGTTTCCGGAGACACAGCTGGAGCAGTGGCCCTTCAAC 240  
QY 241 CTCTTCATGCTGGCGCTTGGTGTGCAGTGGGCAATCTGTGGACGCTTCTTGAGCCAG 300  
DB 241 CTCTTCATGCTGGCGCTTGGTGTGCAGTGGGCAATCTGTGGACGCTTCTTGAGCCAG 300  
QY 301 TTCCCTTCTGGGAAGTGGTCAACACTGTTTCAGTATTCGGTGGCCACCATGAGNGCT 360  
DB 301 TTCCCTTCTGGGAAGTGGTCAACACTGTTTCAGTATTCGGTGGCCACCATGAGNGCT 360  
QY 361 TTGTCGCTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTG 420  
DB 361 TTGTCGCTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTG 420  
QY 421 GTGATGCTGTGCTGGAGGTGACAGCTTTAGGCAACCTGAGGATGGTCATCAGTAATATC 480  
DB 421 GTGATGCTGTGCTGGAGGTGACAGCTTTAGGCAACCTGAGGATGGTCATCAGTAATATC 480  
QY 481 TTCAACACAGACTACCAACATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
DB 481 TTCAACACAGACTACCAACATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
QY 541 CTGCTCTGGGCTGGTCTGCCAAGCTCTACCCGAGGGAACGAGGATTAAGATCAG 600  
DB 541 CTGCTCTGGGCTGGTCTGCCAAGCTCTACCCGAGGGAACGAGGATTAAGATCAG 600  
QY 601 ACAGCAAGCATACCCAGTTGCTGCGCATGCTGGCGGCCCTCTTCTTGTGATGTTCTGG 660  
DB 601 ACAGCAAGCATACCCAGTTGCTGCGCATGCTGGCGGCCCTCTTCTTGTGATGTTCTGG 660  
QY 661 CCAAGTTTCACTCTGCTCTCTGAGAGTCCCAATCGAAAGGAAGATGCCGTGTTCAAC 720  
DB 661 CCAAGTTTCACTCTGCTCTCTGAGAGTCCCAATCGAAAGGAAGATGCCGTGTTCAAC 720  
QY 721 ACCTACTATGCTGTAGCAGTCTGAGGATGAGCAGCATCTCAGGCTCATCTTGGCTCAC 780  
DB 721 ACCTACTATGCTGTAGCAGTCTGAGGATGAGCAGCATCTCAGGCTCATCTTGGCTCAC 780  
QY 781 CCCCAGGGAAGATCAGCAAGCTTATGTGCAAGTGGGCTGTTGGCAGGAGGCGTGGCT 840  
DB 781 CCCCAGGGAAGATCAGCAAGCTTATGTGCAAGTGGGCTGTTGGCAGGAGGCGTGGCT 840

QY 841 GTGGGTACCTCGTGTCCACCTGATCCCTTCTCGTGGCTTGCATGCTGGTCTGTGTG 900  
DB 841 GTGGGTACCTCGTGTCCACCTGATCCCTTCTCGTGGCTTGCATGCTGGTCTGTGTG 900  
QY 901 GCTGGGCTGATCTCCGTGCGGGAGCCAAAGTACCTGCGGGGTGTTGTAACCGAGTCTG 960  
DB 901 GCTGGGCTGATCTCCGTGCGGGAGCCAAAGTACCTGCGGGGTGTTGTAACCGAGTCTG 960  
QY 961 GGGATTTCCCAACAGCTCCATCATGGGCTCAACTTCAGCTTGTCTGGGTCTGTTGGAGAG 1020  
DB 961 GGGATTTCCCAACAGCTCCATCATGGGCTCAACTTCAGCTTGTCTGGGTCTGTTGGAGAG 1020  
QY 1021 ATCATCTACATTTGTGCTGTGTGATACCGTCCGAGCCGGCAATGGCATGATGGC 1080  
DB 1021 ATCAICTACATTTGTGCTGTGTGATACCGTCCGAGCCGGCAATGGCATGATGGC 1080  
QY 1081 TTCCAGGTCCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGTGATAGCTCTCAGCTCT 1140  
DB 1081 TTCCAGGTCCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGTGATAGCTCTCAGCTCT 1140  
QY 1141 GGTCTCTGACAGGTTTGTCTCTAAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200  
DB 1141 GGTCTCTGACAGGTTTGTCTCTAAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200  
QY 1201 TATTTTGTGACCAAGTTTCTGGAAGTTTCTCATTGCTGCTGATGATTTTAA 1254  
DB 1201 TATTTTGTGACCAAGTTTCTGGAAGTTTCTCATTGCTGCTGATGATTTTAA 1254

RESULT 2  
ADE09785

ID ADE09785 standard; DNA; 2811 BP.

XX AC ADE09785;

XX XX 29-JAN-2004 (first entry)

XX Novel DNA-related contig nucleotide sequence #507.

XX novel gene; novel protein; tissue marker; molecular weight marker;  
XX chromosome marker; genetic disorder; contig; ds.

XX Unidentified.

XX WC2003054152-A2.

XX PD 03-JUL-2003.

XX XX 10-DEC-2002; 2002WO-US039555.

XX PR 10-DEC-2001; 2001US-0339739P.

XX PR 11-DEC-2001; 2001US-0339453P.

XX PR 14-MAR-2002; 2002US-0365091P.

XX PR 14-MAR-2002; 2002US-0365384P.

XX PR 12-APR-2002; 2002US-0372381P.

XX PR 22-APR-2002; 2002US-00128558.

XX PR 24-APR-2002; 2002US-0376045P.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QH, Wang J;

XX PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;

XX Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

XX WPI; 2003-569235/53.

XX New polynucleotides, useful for expressing recombinant proteins for  
XX analysis, characterization or therapeutic use, or as markers for tissues  
XX in which the corresponding protein is preferentially expressed.  
XX Disclosure; SEQ ID NO 2329; 1177pp; English.



CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responses for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
XX Sequence 487 BP; 86 A; 120 C; 156 G; 119 T; 0 U; 6 Other;  
SQ  
Query Match 29.2%; Score 365.8; DB 8; Length 487;  
Best Local Similarity 99.5%; Pred. No. 2.7e-95;  
Matches 367; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 706 AATGCCGTGTTCAACACCTACTATGCTGTAGCAGTCAGCGTGTGACAGCCATCTCAGGG 765  
Db 56 ATTCCCGTGTTCACACCTACTATGCTGTAGCAGTCAGCGTGTGACAGCCATCTCAGGG 115  
QY 766 TCATCCTTGCTCACCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGGCGTGTG 825  
Db 116 TCATCCTTGCTCACCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGGCGTGTG 175  
QY 826 GCAGGAGCGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885  
Db 176 GCAGGAGCGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 235  
QY 886 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945  
Db 236 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 295  
QY 946 TGTAAACGAGTGTGCGGATTCCTCCACAGCTCCATCATGCTGCTGCTGCTGCTGCTGCTG 1005  
Db 296 TGTAAACGAGTGTGCGGATTCCTCCACAGCTCCATCATGCTGCTGCTGCTGCTGCTGCTG 355  
QY 1006 GGTCTGCTGGAGAGATCATCTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065  
Db 356 GGTCTGCTGGAGAGATCATCTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 415  
QY 1066 AATGCCATG 1074  
Db 416 AATGCCATG 424  
RESULT 4  
ID ABK12333 standard; cDNA; 1805 BP.  
XX AC ABK12333;  
XX AC ABK12333;  
XX 05-JUN-2002 (first entry)  
XX cDNA encoding human nonerythroid Rh glycoprotein RhBG.  
XX Human; nonerythroid Rh glycoprotein; Rhbg; RhBG; antibody; immunogen;  
KW Rh type B transporter gene; polytypic transporter-type protein;  
KW ion transporter; chromosome 1q21.3; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 39..1415  
FT FT /\*tag= a  
FT FT /product= "Human RhBG protein"  
FT FT 1248..1377  
FT FT /\*tag= b  
FT FT /note= "Encodes C-tail"  
FT FT 1769..1774  
FT FT /\*tag= c  
FT FT /standard\_name= "PolyA signal"

/note= "Atypical polyadenylation site"  
WO200216396-A2.  
28-FEB-2002.  
17-AUG-2001; 2001WO-US025881.  
21-AUG-2000; 2000US-0226767P.  
(NYEL-) NEW YORK BLOOD CENT INC.  
Huang C, Liu Z;  
WPI; 2002-280905/32.  
P-PSDB; AAU78091.  
Homologs of mouse or human nonerythroid Rh glycoproteins, Rhbg or RhBG,  
respectively, useful in the production of antibodies which are useful for  
detecting Rhbg or RhBG glycoproteins in a sample.  
Claim 1; Fig 1a; 59pp; English.  
The present invention relates to a new protein or peptide comprising an  
amino acid sequence having at least 60% identity to a sequence comprising  
455 (mouse nonerythroid Rh glycoprotein homologue (Rhbg)) or 458 (human  
nonerythroid Rh glycoprotein homologue (RhBG)) amino acids, fully defined  
in the specification. The antibody of the invention is useful for  
detecting an Rhbg or an RhBG glycoprotein in a sample, by contacting the  
sample with antibody under conditions suitable for binding, assessing the  
specific binding to the antibody, and thus detecting the presence of an  
epitope of Rhbg or RhBG in the sample. The nucleic acids of the invention  
are useful as probes for detecting transporter genes and particularly Rh  
type B transporter genes including e.g. NH 4<sup>+</sup> ion transporters. The  
present nucleic acid sequence is that of the human RhBG gene located on  
chromosome 1q21.3. This sequence encodes the human RhBG protein of the  
invention. RhBG is a polytypic transporter-type protein  
Sequence 1805 BP; 351 A; 550 C; 515 G; 389 T; 0 U; 0 Other;  
Query Match 14.1%; Score 177; DB 6; Length 1805;  
Best Local Similarity 49.5%; Pred. No. 2.7e-40;  
Matches 456; Conservative 0; Mismatches 465; Indels 0; Gaps 0;  
QY 141 CTATCAAGTTGGCCAGATCTGACCGTGATGGCGGCCAATGGCTTGGCTTCCCTCAGCTC 200  
Db 218 CTACCAAGCTTCCAGACGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 277  
QY 201 GAGTTTCCGGAGACACAGCTGGAGCAGTGGCCCTTCAACCTCTTCATGCTGGCGCTTGG 260  
Db 278 CTTCCTGCAGCGTTACGGCTTCAGCAGCGTGGGCTTCACCTTCTCTGCGCGCTTTTGC 337  
QY 261 TGTGAGTGGCAATCTGCTGCGGAGCGGCTTCCTGAGCAGTTCCTCTCTGGAGAGTGGT 320  
Db 338 CCTGAGTGGTCCACACTGGTCCAGGGCTTTCTCCACTCTTCACGGTGGCCACATCCA 397  
QY 321 CATCACACTGTTCAAGTATTCGGCTGGCCACCATGAGTGCCTTTGCTGGTCTGATCTCAGT 380  
Db 398 TGTGCTGCTGGAGAGCATGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 457  
QY 381 GGATGCTGCTGTTGGGAGAGTCAACTGGCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 440  
Db 458 TGGTGGCTGCTGGGCAAGACCGGCGCTTACCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 517  
QY 441 GACAGCTTTAGCAACCTGAGGATGGTTCATCAGTAATATCTTCAACACAGACTACCAT 500  
Db 518 GGTGCTGTTGGCATCAATGATGTTGCTCTCTTCATCTCTCTGGGGTGAGAGATGCCGG 577  
QY 501 GAACATGATGACATCTACGTTGTCGAGCCCTATTTTGGGCTGTCTGTGGCTGTGTGCTGT 560  
Db 578 AGGCTCCATGATATCCACACCTTTGTTGGCTTACTTTCGGGCTCGTCTCTTTCGGGCTTCT 637  
QY 561 GCCAAGCCTCTACCCGAGGAGACGGAGATAAAGATCAGACAGCAACGATACCCAGTTT 620

Db 638 GTACAGGCCCCAGCTGGAGAGAGACAGCCGCCAGGCTCGCTACCAATTCAGACCT 697  
QY 621 GTCTGCAATCTGGGCGCCCTCTCTTGTGTGATGTTCTGGCCAAAGTTCAACTCTGCTCT 680  
Db 698 CTTGCGCATGATTGGGACCACTCTCTCTGTGTGATCTCTGCGCTAGCTTCAATGCTGCACT 757  
QY 681 GCTGAGAAGTCCCAATCAAGAGGAAGATGCCGTGTTTCAACACACTACTATGCTAGCAGT 740  
Db 758 CACAGCGCTGGGCGTGGGAGCATCGAGCGGCCCTCAACACATACTACTCTCGCTGC 817  
QY 741 CAGCGTGTGACAGCCATCTCAGGTTATCTTGGCTCACCCCAAGGGAAGATCAGCAA 800  
Db 818 CAGCAGCCCTTGGCACCCTTTCCTTGTGAGCCCTTGTAGGGAAGATGGAGGCTTCAAT 877  
QY 801 GACTATGTGACAGTGGCGTGTGGCAGAGCGCTGTGGTGGTACTCTGTGTCACT 860  
Db 878 GTCCCAATCCAAATGCAGCGCTGGCTGGAGGGTGTGTGGGACCTCAAGTGAAT 937  
QY 861 GATCCCTTCTCGCTGGCTTCCCATGTGTGCTGTGCTGTGCTGGCTGATCTCCGTCGG 920  
Db 938 GATGCTGACACCTTGGGCTCTGGCAGTGGCTTCTGGCTGGGACTGTCTCCAGCT 997  
QY 921 GAGGCAAGTACTCCGCGGGTGTGTAAACCGAGTGTGGGATTCGCCACAGTCCAT 980  
Db 998 GGGGTACAAGTCTTCAAGCCCATCTTGAATCAAAATCAAGTCCAAAGTCAAGTGTGG 1057  
QY 981 CATGGGCTCAACTTCACTTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1040  
Db 1058 AGTCCAAACCTCCATGGGATGCGGGGCTCTGCGGGGCTCTGCGGGGCTCTGCGGG 1117  
QY 1041 GGTGCTTGATACCGTCCGAGC 1061  
Db 1118 TGGACTTGCACCCATGAAGC 1138

RESULT 5  
ID ABZ11480 standard; cDNA; 1792 BP.  
XX ABZ11480;  
AC ABZ11480;  
XX 20-JAN-2003 (first entry)  
DE Human polynucleotide SEQ ID NO 362.  
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;  
KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
KW arthritis; cytostatic; immunomodulator; neoplastic; neuroprotective;  
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
KW antiarthritic; gene; ss.  
OS Homo sapiens.  
XX WO200270539-A2.  
PN 12-SEP-2002.  
XX 05-MAR-2002; 2002WO-US0005095.  
XX 05-MAR-2001; 2001US-00799451.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Wehrman T, Wang J, Wang D, Drmanac RT;  
XX WPI; 2002-759812/82.  
DR P-PSDB; ABP69263.

XX PT New polynucleotides comprising sequences assembled from expressed  
PT sequence tags (ESTs), useful for treating cell-proliferative,  
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
PT or coagulation disorders.  
XX  
PS Claim 1; SEQ ID NO 362; 1012pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated polynucleotide (I) comprising a  
CC nucleotide sequence selected from any of 948 sequences (ABZ11119-  
CC ABZ12066) or their mature protein coding portion, active domain coding  
CC protein or complementary sequences. The polynucleotides are useful for  
CC identifying expressed genes or for physical mapping of human genome. The  
CC encoded polypeptides (ABP69849-ABP69849) are useful as molecular weight  
CC markers, as a food supplement, for generating antibodies, in medical  
CC imaging, screening and diagnostic assays and for treating cell-  
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,  
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
CC arthritis, etc. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1792 BP; 340 A; 547 C; 511 G; 394 T; 0 U; 0 Other;

Query Match 14.0%; Score 175.4; DB 6; Length 1792;  
Best Local Similarity 49.4%; Pred. No. 8e-40;  
Matches 455; Conservative 0; Mismatches 466; Indels 0; Gaps 0;  
QY 141 CTATCAAGTTGGCCAGATCTGACCGTATGCGGCCCATTTGGCTTCTCTCACCTC 200  
Db 211 CTACCAAGCTTCCAGGAGCTGATGCCATGGTCTTGTGGGCTTTGACTTCTCATGT 270  
QY 201 GAGTTTCCGGACAGCTGGAGCAGTGTGCGCTTCACTTCTCATGCTGGCGCTGG 260  
Db 271 CTTCTGAGCGTTACGGCTTCAGCAGCGTGGGCTTCACTTCTCTGGCGCTTTCG 330  
QY 261 TGTGAGTGGGCAATCTCTGCTGGAGCGCTTCTGAGCCAGTTCCTTCTGGGAAGTGT 320  
Db 331 CCTGAGTGTCCACACTGGTCCAGGCTTCTTCACTTCTTCCAGGTTGGCCATCA 390  
QY 321 CATCACTGTTTCAATTTGCGCTGGCCACCATGATGCTTTGTGCGTCTGATCTCAGT 380  
Db 391 TGTGCGCTGGAGAGCATGATCAATGCTGACTTTTGTGCGGGGCGCTCATCTCTCT 450  
QY 381 GGATCTCTCTTGGGGAAGTCACTTGGCGAGTGTGGTGGTGGTGGTGGTGGTGGT 440  
Db 451 TGGTCCCTCTCTGGGCAAGACCGGGCTTACCCAGCTGCTGCTCATGCGCTTCTGGAGT 510  
QY 441 GACAGCTTTAGGCAACCTTGAGGATGCTCATCAGTAATATCTTCAACACAGACTACCAT 500  
Db 511 GGTGCTGTTGGCATCAATGATTTGTCTCTTCACTTCTCTGGGGTGGAGATGCGCG 570  
QY 501 GAACATGATGACATCTACGTGTTGCGAGCCTATTTTGGGCTGTCTGTGGCTGTGCT 560  
Db 571 AGGCTCCATGACTATCCACACTTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 630  
QY 561 GCCAAAGCTCTACCCGAGGAGCGGAGGATTAAGATCAGACAGCAAGATACCATTT 620  
Db 631 GTACAGCCCCAGCTGGAGAAAGCAAGCACCGCGAGGCTCCGCTTACCAATCAGACCT 690  
QY 621 GTCTGCCATGCTGGGCGCCCTCTTCTTGTGGATGTTCTGGCAAGTTCACACTCTGCT 680  
Db 691 CTTGCGCATGATTGGGACCACTTCTTGTGATCTTCTGGCTAGCTTCAATGCTGACT 750  
QY 681 GCTGAGAAGTCCAAATCGAAGGAAGATGCGTGTTCACACCTTACTATGCTGTAGCAGT 740  
Db 751 CACAGCGTGGGGCTGGGAGCATCGGAGCGCCCTCAACACATACTTCTCTGGCTGC 810  
QY 741 CAGCGTGTGACAGCCATCTCAGGCTCATCTTGGCTCACCCCCAAGGAAGATCAGCAA 800

Db 811 CAGACCCCTTGGACACTTTGCTTGTACCCCTTGTAGGGAGATGGGAGCTTGACAT 870  
QY 801 GACTTATGTGCACAGTGGCGTGTGTCACAGAGCGTGGCTGGGTAACCTGCTGTCACT 860  
Db 871 GGTCCACATCCAAATGACAGCGTGGCTGGAGGGTGTGGTGGGACCTCAAGTGAAT 930  
QY 861 GATCCCTTCTCGTGGCTGCGCATGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 920  
Db 931 GATGCTGACACCTTTGGGGCTCTGGCAGCTGGCTTCTTGGCTGGGACTGTCTCCAGCT 990  
QY 921 GGGAGCAAGTACTGCGGGGTGTGTAAACCGAGTGTGGGATTCGCCACAGCTCCAT 980  
Db 991 GGGGTACAGTCTTTCACGCCCATCTTGAATCAAAATTCAGAGTCCAGACACATGTGG 1050  
QY 981 CATGGGCTCAACTTCAGCTTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1040  
Db 1051 AGTCCACAACTCCATGGATGCGGGGTCTCTGGGGGCGCTCTGGGGGTCTTGTGGC 1110  
QY 1041 GGTGCTGTACCTGGGAGC 1061  
Db 1111 TGGACTTGGACCCCATGAAGC 1131

RESULT 6  
ID ACD20392/c  
XX ACD20392 standard; DNA; 1351 BP.  
AC ACD20392;  
XX  
DT 26-AUG-2003 (first entry)  
XX  
DE DNA encoding human NOV15 protein.  
XX  
KW Human; NOVX; inflammatory disorder; demyelination disease; stroke;  
KW renal disorder; infection; cardiomyopathy; atherosclerosis; acne;  
KW hypertension; pancreatitis; Von Hippel-Lindau; endometriosis; fertility;  
KW scleroderma; cirrhosis; inflammatory bowel disease; Crohn's disease;  
KW hemophilia; autoimmune disease; allergy; AIDS;  
KW graft versus host disease; Alzheimer's disease; arthritis; pain;  
KW Parkinson's disease; Huntington's disease; obesity; diabetes;  
KW hair growth; hair loss; asthma; schizophrenia; glomerulonephritis;  
KW lupus erythematosus; psoriasis; antidiabetic; anorectic; metabolic;  
KW neutropenic; neuroprotective; cytostatic; antibacterial; virucide;  
KW proteoacid; antiarteriosclerotic; hypotensive; cerebroprotective;  
KW antiinflammatory; gynaecological; antiinfertility; dermatological;  
KW hepatotropic; haemostatic; immunosuppressive; antiallergic;  
KW antiarthritic; anticonvulsant; antiseborrhoeic; antiasthmatic;  
KW neuroleptic; anti-HIV; analgesic; nephrotropic; antipsoriatic; gene; ds.  
XX Homo sapiens.  
XX  
XX WO200298917-A2.  
XX  
XX PD 12-DEC-2002.  
XX  
XX PF 12-FEB-2002; 2002WO-US022049.  
XX  
XX 12-FEB-2001; 2001US-0268221P.  
XX 13-FEB-2001; 2001US-0268496P.  
XX 14-FEB-2001; 2001US-0268646P.  
XX 14-FEB-2001; 2001US-0268665P.  
XX 15-FEB-2001; 2001US-0269136P.  
XX 16-FEB-2001; 2001US-0269310P.  
XX 16-FEB-2001; 2001US-0269530P.  
XX 15-MAR-2001; 2001US-0276403P.  
XX 16-MAR-2001; 2001US-0276399P.  
XX 16-MAR-2001; 2001US-0276703P.  
XX 23-MAR-2001; 2001US-0278199P.  
XX 28-MAR-2001; 2001US-0279274P.  
XX 30-MAR-2001; 2001US-0280238P.  
XX 02-APR-2001; 2001US-0280899P.  
XX 08-AUG-2001; 2001US-0310797P.  
XX 14-AUG-2001; 2001US-0312284P.

PR 14-SEP-2001; 2001US-0322294P.  
PR 14-SEP-2001; 2001US-0322295P.  
PR 18-OCT-2001; 2001US-0330293P.  
PR 31-OCT-2001; 2001US-033104P.  
PR 31-OCT-2001; 2001US-033109P.  
PR 21-NOV-2001; 2001US-033127P.  
PR 28-NOV-2001; 2001US-0331772P.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Guo X, Fernandes E, Li L, Kekuda R, Liu Y, Leite M, Spytek KA;  
XX Ji W, Casman SJ, Boldog FL, Paturajan M, Vernet CM, Ballinger RA;  
XX Malyankar UM, Tchernev VT, Bialock AD, Gusev V, Rastelli L;  
XX Mezes PD, Ellerman K, Heyes M, Herrmann JL, Shinkets RA, Ioime N;  
XX Pena CEA, Shenoy SG, Taupier RJ, Gerlach V, Gorman L;  
XX  
XX WPI; 2003-148650/14.  
XX P-PSDB; ABO15004.  
XX  
XX Novel NOVX polypeptide useful for identifying an agent that binds to the  
XX polypeptide, and for treating cardiomyopathy, atherosclerosis,  
XX hypertension, infertility, scleroderma, cirrhosis, and inflammatory bowel  
XX disease.  
XX  
XX Claim 9; Page 115; 566pp; English.  
XX  
XX The present invention relates to the isolation of novel human  
XX polypeptides referred to as NOVX (NOV1-NOV37), variants of these  
XX proteins, and the polynucleotide sequences encoding them. The NOVX  
XX proteins of the invention share homology to various types of protein  
XX families such as zinc finger-like proteins, enzymes, receptors, and  
XX lipoproteins. The sequences of the invention may be useful in the  
XX manufacture of a medicament for treating a syndrome associated with a  
XX human disease. For example they can be used to treat inflammatory  
XX disorders, demyelination disease, renal disorders, infections,  
XX cardiomyopathy, atherosclerosis, hypertension, stroke, pancreatitis, Von  
XX Hippel-Lindau, endometriosis, fertility, scleroderma, cirrhosis,  
XX inflammatory bowel disease, Crohn's disease, haemophilia, autoimmune  
XX diseases, allergies, graft versus host disease, Alzheimer's disease,  
XX arthritis, Parkinson's disease, Huntington's disease, obesity, diabetes,  
XX acne, hair growth/loss, asthma, schizophrenia, AIDS, pain,  
XX glomerulonephritis, lupus erythematosus, and psoriasis. ACD20372-ACD20427  
XX represent DNA sequences encoding the NOVX polypeptides of the invention.  
XX Note: SEQ ID Nos 113-460 are known sequences used for homology purposes  
XX  
XX Sequence 1351 BP; 325 A; 372 C; 389 G; 265 T; 0 U; 0 Other;  
XX  
XX Query Match 13.2%; Score 165.4; DB 7; Length 1351;  
XX Best Local Similarity 49.1%; Pred. No. 5.4e-37;  
XX Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;  
QY 141 CTATCAAGTTGGCCAGATCTGACCGTGTGGCGCCATTGGCTTGGCTTCCCTCAGCTC 200  
Db 1175 CTACCAAGCTTCCAGGACGTGACGTGATGCTTCTGGGCTTGGCTTCCCTCATGAC 1116  
QY 201 GAGTTTCGGAGACACAGCTGGAGACGTGGCCCTTCAACCTCTCATGTGGGCTTGG 260  
Db 1115 TTTCTGCAGCGCTACGGCTTCAGCGCGCTTCAACTTCTGTGGCAGCTTCGG 1056  
QY 261 TGTGACGTGGCAATCTGCTGGAGCGCTTCTCGAGCGATTCCCTTCTGGGAAGTGGT 320  
Db 1055 CATCAGTGGCGCTGCTCATGACAGGCTGGTTCATCTTCAAGACCGCTACATCGT 996  
QY 321 CATCACACTGTTTCAAGTATTCGGCTGGCCACCATGAGTCTTGTGGTGTGATCTCAGT 380  
Db 995 CGTGGCGTGGAGAACCTCATCAACGCTGACTTCTGGGTGGCTTGTCTGGTGGCTT 936  
QY 381 GGATGCTCTCTGGGGAAGTCACTGGCCAGTGTGGTGTGATGTGGTGGTGGT 440  
Db 935 TGGGCGAGTCTTGGGTAAGTCAAGCCCATTCAGCTGTCTCATGATCTTCTTCCAGT 876  
QY 441 GACAGCTTTAGCAACCTGAGGTGGTTCATCAGTAATATCTTCAACACAGATCACCAT 500



Db 875 GACCCTCTCGTGTGAATGAGTTCAATCTCTTAACTGCTAAAGGTGAAGATCAGG 816  
QY 501 GACATGATGACATCTACATGTTTCGACGCTATTTGGCTGTCTGTGGCTTGGTGCCT 560  
Db 815 AGCTCATGACCATCCACATTTGGCGCTACTTTGGGCTCACAGTGACCCGGATCCT 756  
QY 561 GCCAAAGCTCTTACCAGGAGACGGAGGATAAAGATCAGACAGCAACGATACCCAGTTT 620  
Db 755 CTACCGAGCAACTAGACAGAGCAGAGGAGACAGATTTCTGTACCATGTCGACCT 696  
QY 621 GTCTGCATGCTGGGCGCCTCTCTTCTGTGGATGTTCTTGCCCAAGTTTCAACTCTGCTCT 690  
Db 695 CTTTGGCATGATTGGCACCCTCTCTCTGTGATGTACTGGCCCGAGCTTCAACTCAGCAT 636  
QY 681 GCTGAGAAGTCCNAATCAAGAGGAGATGCCGTGTTCAACACTACTATGCTGTAGAGT 740  
Db 635 ATCCTACCATGGGACAGCCAGCACCGAGCGCCCATCAACACTACTGCTCTTGGCAGC 576  
QY 741 CAGCTGTGTGACAGCCATCTCAGGGTCACTCTTGGTCACTCCGCCCAAGGGAAGATCAGCA 800  
Db 575 CTGCGTGTACTCTCGTGGCATATCCAGTGCCTGCAAGAGAGGCAAGCTGGACAT 516  
QY 801 GACTTATGTGACAGTGGGTGTGGCAGAGCGGTGGCTGTGGGTACCTCTGTGTCACCT 860  
Db 515 GGTGCATATCCAGATGCCACGCTCGCAGAGGGGTGGCGTGGGTACCGTGTCTGAGAT 456  
QY 861 GATCCCTCTCGTGGCTTGCATGCTGTGGTCTGTGGTGTGGCTGATCTCGTGG 920  
Db 455 GATGCTCATGCTTACGGTGGCCCTCATCFCGGCTTCGTCTGCGGCATCATCTCCACCCT 396  
QY 921 GGGAGCAAGTACCTCGCGGGTGTGTAACCGAGTGTGGGATTCCTCCACAGCTCCAT 980  
Db 395 GGGTTTGTATACCTGACCCCATCTCTGAGTCCCGCTGCATCCAGACACATGTGG 336  
QY 981 CATGGGCTACAATTGAGTGTGGTGTGGTGTGGTGTGGGATCATCTACATTTG 1035  
Db 335 CATTACAATCTGATGCGATTCCTGGCATCATAGGCGGATCGTGGGTGCTGTG 281

RESULT 7  
AAC64286  
ID AAC64286 standard; cDNA; 1853 BP.  
AC AAC64286;  
XX  
DT 23-FEB-2001 (first entry)  
XX  
DE Human membrane-associated protein HUMAP-13 cDNA.  
XX  
KW Human membrane-associated protein; HUMAP; transgenic organism;  
KW drug screening; cell signalling modulator; agonist; antagonist;  
KW cell differentiation modulator; cell proliferation modulator;  
KW cell proliferative disorder; cancer; cell differentiation disorder;  
KW developmental disorder; cell signalling disorders; endocrine disorder;  
KW hyperpituitarism; hypothyroidism; hyperparathyroidism; infection;  
KW pancreatic disorder; diabetes mellitus; immunological disorder;  
KW hereditary neuropathy; gonadal steroid hormone associated disorder;  
KW infertility; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200065054-A2.  
XX  
PD 02-NOV-2000.  
XX  
PF 20-APR-2000; 2000MO-US010884.  
XX  
PR 23-APR-1999; 99US-0130694P.  
PR 23-JUN-1999; 99US-0140580P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Hillman JL, Bandman O, Tang YT, Lal P, Yue H, Reddy R;

PI Azimzai Y, Baughn MR;  
XX WPI; 2000-687346/67.  
DR P-PSDB; AAB29656.  
XX  
PT Human membrane-associated protein, useful for diagnosis and treatment of  
PT cell signaling, cell differentiation and cell proliferation disorders  
PT such as cancer, and for identifying agonists and antagonists.  
XX  
PS Claim 4; Page 96; 99pp; English.  
XX  
XX The invention relates to 17 human membrane-associated proteins, HUMAP-1  
CC to HUMAP-17 (AA22644-B29660) and the cDNAs encoding them (AAC64274-  
CC C64290). The invention also relates to expression constructs, host cells  
CC and transgenic organisms comprising a HUMAP nucleic acid sequence; the  
CC recombinant preparation of a HUMAP; methods of screening compounds for  
CC their ability to modulate HUMAP activity or expression; and  
CC pharmaceutical compositions comprising a HUMAP protein, a HUMAP agonist  
CC or HUMAP antagonist. The HUMAPs acts as modulators of cell signalling,  
CC differentiation and proliferation. A HUMAP is useful for screening a  
CC compound for effectiveness as an agonist or antagonist of HUMAP activity.  
CC The protein, or the identified agonist or antagonist is useful for  
CC treating a disease or condition associated with decreased or increased  
CC expression of functional HUMAP. A HUMAP nucleic acid is useful for  
CC screening a compound for its ability to alter expression of that  
CC particular HUMAP gene. A wide variety of disease may be treated using  
CC compositions of the invention. These diseases include cell proliferative  
CC disorders (e.g., actinic keratosis, arteriosclerosis); cancer (e.g.,  
CC breast, bladder, bone marrow, brain and uterine cancer); cell  
CC differentiation disorders, in particular developmental disorders (e.g.,  
CC renal tubular acidosis, anaemia, Cushing's syndrome, achondroplasia,  
CC epilepsy, and muscular dystrophy); cell signalling disorders, in  
CC particular endocrine disorders such as hypothalamus and pituitary  
CC disorders resulting from lesions such as thymosis; disorders associated  
CC with hyperpituitarism (e.g., acromegaly); disorders associated with  
CC hypothyroidism (e.g., goitre); hyperparathyroidism; pancreatic disorders  
CC such as type I or type II diabetes mellitus; infections; immunological  
CC disorders; hereditary neuropathies (e.g., neurofibromatosis); and  
CC disorders associated with gonadal steroid hormones (e.g., infertility,  
CC endometriosis, polycystic ovary syndrome, osteoporosis, Leydig cell  
CC deficiency and gynecomastia). Antibodies which specifically bind HUMAP  
CC may be used for the diagnosis of disorders associated with the expression  
CC of HUMAP, or in assays to monitor patients being treated with HUMAP or  
CC agonists, antagonists or inhibitors of HUMAP. The present sequence  
CC represents a HUMAP cDNA of the invention  
XX  
SQ Sequence 1853 BP; 367 A; 563 C; 508 G; 415 T; 0 U; 0 Other;  
Query Match 13.2%; Score 165.4; DB 3; Length 1853;  
Best Local Similarity 49.1%; Pred. No. 6.4e-37;  
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;  
QY 141 CTATCAAGTTGGCCCAAGATCTGACCGTGTATGGCGCCATTGGCTTGGCTTCCACCTC 200  
Db 207 CTACCAAGCTTCAGGACGTGCATGATGGTCTCTGGGCTTGGCTTCCATGAC 266  
QY 201 GAGTTTCGGAGACACAGCTGGAGCAGTGTGGCTTCAACCTTTCATGCTGGCGCTGG 260  
Db 267 TTTCTCTGAGCGCTACGGCTTCAGCGCGTGGGCTTCAACTTCTGTTGGCAGCTCGG 326  
QY 261 TGTGAGTGGGCAATCTCTGAGCGGCTTCTGAGCAGTTCCTTCTGGAGAGGTGGT 320  
Db 327 CATCCAGTGGGCGCTGCTCATGAGGGCTGGTTCCACTTCTTACAAGACCGTACATCGT 386  
QY 321 CATCACACTGTTTCACTATTGCGCTGGCCACCATGAGTGTCTTGTGGTGTGATCTCAGT 380  
Db 387 CGTGGCGTGGAGAACCTCATCAACGTGACTTCTGCGTGGCGCTCTGTCTCGTGGCCTT 446  
QY 381 GGATGCTCTCTTGGGAGAGTCAACTTGGCGCAGTTGGTGGTGTGGTGTGGTGGAGGT 440  
Db 447 TGGGCGAGTTCTGGGTAAGTCAGCCCATTCAGCTGCTCATCATGACTTCTTCCAGT 506  
QY 441 GACAGCTTTAGCAACCTGAGGATGGTTCATCAGTAATATCTTCAACACAGACTACCAT 500

```

507 GACCTCTTCGCTGTAATGAGTTCAATCTCTTAACTGCTAAAGGTGAGATGCAGG 566
501 GAACATGATGACATCTACCTGCTGTCAGCCTATTTTGGGCTGCTGTGGCCCTGGTGCCT 560
567 AGGCTCCATGACCATCCACATTTGGCGCTACTTTGGGCTCACAGTGACCCGATCCT 626
561 GCCAAGCCTCTACCCGAGGAGGAGGTAAGATACAGACACGACGATACCCAGTTT 620
627 CTACCGACGCAACTAGACGAGCAAGGAGAGACAGAAATTCGTGTACAGTCCGACCT 686
621 GTCTGCATCTCGGGCGCCCTCTTTCTGTGATGTTCTGGCCAAAGTTTCAACTCTGTCT 680
687 CTTTGGCATGATGGCACCTCTTCTCTGTGATGTTACTGGCCAGCTTCACTCAGCCAT 746
681 GCTGAGAAGTCCAAATCGAAAGGAAGATGCCGTGTTCAACACCTACTATGCTGTAGAGT 740
747 ATCCTACCATGGGACAGCCAGCACCGAGCGGCATCAACACCTACTGTCTTGGCAGC 806
741 CAGGCTGGTGACAGCCATCTCAGGTCATCTTGGCTCACCCCAAGGGAAGATCAGCA 800
807 CTGCTGTCTTACCTCGGTGCAATATCCAGTCCCTGCAACAAGGGCAAGCTGACAT 866
801 GACTTATGTGACAGTCTCGGTGTTGGCAGAGGCGTGGCTGTGGTACCTCGTGTACCT 860
867 GGTGCACATCCAGATGCCAGCTCGCAGAGGGGTGGCGTGGTACCGCTGCTGAGAT 926
861 GATCCCTTCTCGGTGCTTGCATGCTGCTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 920
927 GATGCTCATGCTTACGCTGCTTACGCTGCTTACGCTGCTTACGCTGCTTACGCT 986
921 GGGAGCAAGTACTCGCGGGGTGTTGTAACCGAGTCTGGGATTCGCCACAGCTCCAT 980
987 GGGTTTGTATACCTTGACCCCAATTCCTGGAGTCCCGGCTGCACATCCAGACATGTTG 1046
981 CATGGGTACAACCTCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1035
1047 CATTAACAATGATGCATCTCTGGCATCATAGCGGCATCGTGGTGTCTGTG 1101

```

RESULT 8  
 AAH25763  
 ID AAH25763 standard; cDNA; 1949 BP.  
 XX  
 AC AAH25763;

14-AUG-2001 (first entry)  
 Oesophagus cancer associated DRC2 coding sequence.  
 DRC2; oesophagus cancer; ss.  
 Unidentified.

Key Location/Qualifiers  
 CDS 41..1480  
 /\*tag= a  
 /product= "DRC2"

CN1283694-A.  
 14-FEB-2001.  
 10-AUG-1999; 99CN-00117523.  
 10-AUG-1999; 99CN-00117523.  
 (ONCO-) ONCOLOGY INST TUMOR HOSPITAL CHINESE ACA.  
 Wang M, Xu Z, Xu X;  
 WPI; 2001-291757/31.  
 P-PSDB; AAB97000.

XX  
 PT  
 XX  
 PS  
 XX  
 CC  
 CC  
 CC  
 CC  
 CC  
 CC  
 CC  
 XX  
 SQ

Gene associated with esophagus cancer.  
 Claim 4; Fig 1; 33pp; Chinese.  
 The present invention provides the protein and coding sequences of DRC2, which is associated with oesophagus cancer. Also provided is a method for diagnosing diseases associated with an abnormal version of the nucleic acid and encoded protein, involving detecting any nucleic acid sequence mutations, methylation and variation at the RNA and polypeptide level. The sequences can be used in the treatment of cancer

Sequence 1949 BP; 382 A; 593 C; 530 G; 444 T; 0 U; 0 Other;

Query Match 13.2%; Score 165.4; DB 5; Length 1949;  
 Best Local Similarity 49.1%; Pred. No. 6.6e-37;  
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCAAAGATCTGACCGTGTATGGCGCCATTTGGCTTGGCTTCTCCTC 200  
 DB 217 CTATCCCAAGCTTCCAGGACGTGCACTGATGGTCTTCTGGGCTTCGGCTTCTCATGAC 276  
 QY 201 GAGTTTCCGAGACACAGCTGGAGAGTGTGGCTTCAACCTCTTCACTGCTGGCTTGG 260  
 DB 277 TTTCTGCGAGCGCTACGGCTTCAAGCCGCTGGCTTCAACTTCTGTGGCAGCTTCGG 336  
 QY 261 TGTGAGTGGGCAATCTCTGTCAGCGCTTCTGAGCCAGTTCCTTCTGGGAAGTGGT 320  
 DB 337 CATCCAGTGGGCGCTGCTCATGACAGGCTGGTTCACCTTTACAGACCGCTACATCGT 396  
 QY 321 CATCACACTGTTTCAAGTATTCGGCTGGCCACCATGAGTGTCTTGTGGTGTGATCTCAGT 380  
 DB 397 CGTGGCGTGGAGAACCTCATCAACGTGACTTCTCGTGGCGCTCTGTCTGGCTGGCCTT 456  
 QY 381 GAGTGTCTTGGGAGAGTCACTTGGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 440  
 DB 457 TGGGGCAGTCTTGGGTAAGTCAAGCCCAATTCAGCTGCTCATGATGACTTCTTCCAAGT 516  
 QY 441 GACAGCTTTAGGCAACCTGAGGATGGTCACTAGTAATATCTTCAACACAGACTACACAT 500  
 DB 517 GACCTCTTCGTGTGATGAGTTTCACTTCTTAACTGCTAAAGTGAAGATGCAGG 576  
 QY 501 GAACATGATGACATCTACCTGTTGCGAGCTATTTTGGGCTGTCTGTGGCTGTGGTGCCT 560  
 DB 577 AGCTCCATGACCATCCACATTTGGCGCTACTTTGGGCTCACAGTACCCGATCCT 636  
 QY 561 GCCAAGCCTCTACCCGAGGAGCGGAGGATGAAGTACAGACGACGATACCCAGTTT 620  
 DB 637 CTACCGACGCAACTAGACGAGACGAGAGACAGAAATTCGTGTACAGTCCGACCT 696  
 QY 621 GTCTGCCATGCTGGGCGCCCTCTTCTGTGATGTTCTGGCCAAAGTTTCAACTCTGTCT 680  
 DB 697 CTTTGGCATGATGGCACCTCTTCTGTGATGTTACTTGGCCAGCTTCACTCAGCCAT 756  
 QY 681 GCTGAGAAGTCCAAATCGAAAGGAAGATGCCGTGTTCAACACCTACTATGCTGTAGAGT 740  
 DB 757 ATCTTACCTAGGAGACAGCCAGCCAGCGCCATCAACACCTACTTGTCTTGGCAGC 816  
 QY 741 CAGGCTGGTGACAGCCATCTCAGGTCATCTTGGCTCACCCCAAGGGAAGATCAGCAA 800  
 DB 817 CTGCTGCTTACCTCGGTGCAATATCCAGTCCCTGCAACAAGAGGGAGCTGGACAT 876  
 QY 801 GACTTATGTGACAGTCTGGGTGTTGGCAGGAGGCGTGGCTGTGGTACCTCTGTGTACCT 860  
 DB 877 GGTGCACATCCAGATGCCAGCTCGCAGAGGGGTGGCGTGGTACCGCTGTGTAGAT 936  
 QY 861 GATCCCTTCTCGGTGCTTGCATGCTGTGGTGTCTTGTGGTGGGCTGATCTCCGTCGG 920  
 DB 937 GATGCTCATGCTTACGCTGCTTACGCTGCTTACGCTGCTTACGCTGCTTACGCT 996  
 QY 921 GGGAGCAAGTACTCGCGGGGTGTTGTAACCGAGTCTGGGATTCGCCACAGCTCCAT 980  
 DB 997 GGGTTTGTATACCTTGACCCCAATTCCTGGAGTCCCGGCTGCACATCCAGGACATGTTG 1056

QY 981 CATGGCTACAACTTCAGTCTGCTGGTCTGCTGGAGAGATCACTACATGTG 1035  
 Db 1057 CATTAAACAATCTGCATGGCAATTCCTGGCATCATAGGCGCATCGTGGTGTGTG 1111

## RESULT 9

ABK49215  
 ID ABK49215 standard; cDNA; 1952 BP.

XX AC ABK49215;

XX AC  
 DT 15-JUL-2002 (first entry)

XX DE cDNA encoding human Rh type C gene (RHCG) protein.

XX RHCG: human; non-erythroid Rh type C glycoprotein; chromosome 15q25;  
 KW gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 24..1440

FT /tag= a

FT /product= "RhCG protein"

FT polyA\_signal 1985..1990

FT /tag= b

XX WO200220719-A2.

XX PN 14-MAR-2002.

XX PD 05-SEP-2001; 2001WO-US027503.

XX PF 07-SEP-2000; 2000US-0230660P.

XX PR (NYBL-) NEW YORK BLOOD CENT INC.

XX PA Huang C, Liu Z;

XX PI WPI; 2002-351774/38.

XX DR P-PSDB; AAU78997.

XX PT Nucleic acid sequences encoding novel mammalian nonerythroid Rh type C  
 and glycoproteins which have a characteristic twelve transmembrane domain  
 structure.

XX PS Claim 2; Fig 1; 53pp; English.

XX CC This invention relates to the nucleic acid and protein sequences of novel  
 human and mouse non-erythroid Rh type C glycoprotein (RhCG). The RhCG  
 protein and the mouse homologue (rhcg) have a characteristic 12  
 transmembrane domain structure and are expressed in kidneys and testis.  
 The invention also comprises a method for antibody that specifically  
 binds an epitope of the glycoprotein and a method for detecting the  
 protein using this antibody. The antibodies of the invention may be used  
 in Western blots, enzyme linked immunosorbent assays (ELISA) or  
 immunohistochemical assays to identify the non-erythroid tissues,  
 particularly kidney and testis, that express the RhCG or Rhcg  
 glycoproteins. The methods are used for detecting an Rhcg or and RhCG  
 glycoprotein in a sample. The present sequence represents the cDNA  
 encoding the human RhCG Rh type C glycoprotein (RhCG) protein sequence of  
 the invention. The gene encoding this protein is located on human  
 chromosome 15q25

XX SQ Sequence 1952 BP; 398 A; 586 C; 524 G; 444 T; 0 U; 0 Other;

XX Query Match

XX Best Local Similarity 13.2%; Score 165.4; DB 6; Length 1952;

XX Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGCCAAGATCTGACCGTGATGGCGGCATTTGGGTCTCTCACCTC 200  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 201 CTATCCAAAGCTTCAGGAGTGTCACGTGATGTCTCTGGTGGGCTTCGGGCTTCTCTATGAC 260  
 QY 201 GAGTTTCGGGAGACACAGCTGGAGCAGTGTGGCCCTTCAACCTCTTTCATGCTGGCGTTGG 260  
 Db 261 TTTCCTGCAGCGCTACGGCTTCAGCGCGGTGGGCTTCAACTTCTCTGTTGGCAGCCTTCGG 320  
 QY 261 TGTGCAAGTGGCAATCTCTGCTGGAGCGGCTTCTGAGCCAGTTCCTCTTGGGAAGGTGTG 320  
 Db 321 CATCCAGTGGCGCTGCTCATGACAGGCTGGTTCACCTTCTTACAGACCGGTACATCGT 380  
 QY 321 CATCACACTGTTCACTATTTCGGCTGCCACCATGATGCTGTTGTGGTGGTCTGATCTCACT 380  
 Db 381 CGTGGCGGTGGAGAACCTCATCAAGCTGACCTTCTGGTGGGCTCTGTCTGGTGGGCTT 440  
 QY 381 GGATGCTGTCTGGGAAGGTCAACTTGGCGCAGTTGGTGGTGGTGGTGGTGGTGGTGGTGG 440  
 Db 441 TGGGGCAGTCTGGGTAAGTCAGCCCCATTTCAGCTGCTCATCATGATCTTCTTCTTCAAGT 500  
 QY 441 GACAGCTTTAGCAACCTGAGGATGGTTCATCATGTAATATCTTCAACACAGACTACCAT 500  
 Db 501 GACCTCTTCGCTGTGAATGATTCATTCTCTTAACCTGTAAAGGTGAAGGATGTCAGG 560  
 QY 501 GAACATGATGACATCTACGTGTTTCGACGCTTATTTTGGGCTGTCTGTGGCTGTGGCT 560  
 Db 561 AGGCTCCATGACCATCCACATTTGGGCGCTTCTTTGGGCTCACAGTACCCGATCTCT 620  
 QY 561 GCCAAAGCTCTTACCCGAGGGAACGAGGATAAAGATCAGACAGACAGATACCCAGTTT 620  
 Db 621 CTACCGACGCAACCTTAGAGCAGAGCAGAGACAGAAATCTGTGTACCATGTCGACCT 680  
 QY 621 GTCTGCCATGTGGGGCGCTCTTCTTGTGATGTTCTGGCCAGTTTCAACTCTGCTCT 680  
 Db 681 CTTTGGCATGATTTGGCACCTCTTCTGTGGATGTTCTGGGCGCTTCACTGAGCCAT 740  
 QY 681 GCTGAGAAAGTCCAAATCGAAAGGAAGATGCGGTGTTTCAACACTTACTATGCTGTAGCAGT 740  
 Db 741 ATCTTACCATGGGACAGCCAGCAGCCGACCCCATCACCTACTGCTCTTGGCAGC 800  
 QY 741 CAGCGTGTGACAGCCATCTTCAGGGTCACTCTTGGCTTCACTCCCCAAGGGAAGATCAGCAA 800  
 Db 801 CTGCGTGTCTTACCTCGGTGGCAATATCCAGTCCCTGCACAAAGAGGCAAGCTGGACAT 860  
 QY 801 GACTTATGTGACAGTGGGTGTTGGCAGAGGCGTGGGTGAGTCTGCTGTGACCT 860  
 Db 861 GGTGCATATCCAGATGCCAGCTCGCAGAGGGGTGGCCGCTGAGTCTGCTGAGAT 920  
 QY 861 GATCCCTTCTCGGTGGCTTGCATGGTGTGCTGCTGGGCTGATCTCGTCTGG 920  
 Db 921 GATGCTCATGCTTACGGTGGCTTCTGCTGGGCTTCTGCTGGGCTCATCTCCACCT 980  
 QY 921 GGGAGCCCAAGTACCTGCGGGGTGTTGTAACGAGTGTCTGGGGATTCCTCCACAGTCCAT 980  
 Db 981 GGGTTTGTATACCTGACCCCATTCCTGAGTCCCGGTGCACATCCAGGACACATGTGG 1040  
 QY 981 CATGGGCTTACAACTTCAGTCTGCTGGTCTGTTGGAGAGATCATCTACATGTG 1035  
 Db 1041 CATTAAACAATCTGCATGGCATTCCTGGCATCATAGGCGGCATCGTGGTGTGCTGTG 1095

## RESULT 10

ABX76257

ID ABX76257 standard; DNA; 1968 BP.

XX AC ABX76257;

XX DT 02-APR-2003 (first entry)

XX DE Lung cancer-associated polynucleotide #123.

XX KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;  
 antinflammatory; antistatic; non-small cell lung cancer; atelectasis;  
 small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
 chronic obstructive pulmonary disease; hypersensitivity pneumonitis;



XX PD 05-OCT-2000.  
 XX PF 31-MAR-2000; 2000WO-US008621.  
 XX PR 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 XX 30-MAR-2000; 2000US-00540763.  
 PR (CURA-) CURAGEN CORP.  
 PA Shimkrets RA, Leach M;  
 XX WPI: 2000-602362/57.  
 XX P-PSDB; AAB42135.  
 XX Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX Claim 5; Page 2946-2948; 5507pp; English.  
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
 CC anticonvulsant; antiarthritic; immunosuppressive; immunostimulant;  
 CC cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antineoplastic; antitumor; antibacterial;  
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX Sequence 1970 BP; 400 A; 593 C; 530 G; 445 T; 0 U; 2 Other;  
 SQ

Query Match 13.2%; Score 165.4; DB 3; Length 1970;  
 Best Local Similarity 49.1%; Pred. No. 6.e-37;  
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCCAAGATCTGACCGTGTATGGCGCCCATTTGGCTTCCCTCCTCCTGATGAC 278  
 Db 219 CTACCAAGCTTCCAGGACGTGACGCTGATGCTTCTCGTGGGCTTCGGCTTCCCTCCTGATGAC 278

QY 201 GAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAACCTTCTCATGCTGGGCTTGG 260  
 Db 279 TTCTCTGACGCTACGCTTACGCGCGTGGCTTCACTTCTGTTGGCAGCTTCGG 338

QY 261 TGTGAGTGGGCAATCTCTGCTGGAGCGTCTCTGAGCCAGTTCCTTCTGGGAAGGTGGT 320  
 Db 339 CATCCAGTGGGCGTGTCTATGACGCGTGGTTCACCTTCTTACAAGCCGCTACATCGT 398

QY 321 CATCACACTGTTCAGTATTCGGCTGGCCACCATGATGCTTGTTCGGTGTGATCTCAGT 380  
 Db 399 CGTGGGCGTGAGAACCTTCATCAACGCTGACTTCTGCGTGGCCCTCTGCTCGGTGGCCTT 458

QY 381 GGATGCTCTCTTGGGGAAGGTCAACTTGGCGCAGTGTGGTGTGATGCTGCTGGTGGAGGT 440  
 Db 459 TGGGSCAGTTCTGGGTAAGTCAGCCCATTCAGCTGCTCATCATGACTTCTTCCAGT 518

QY 441 GACAGCTTTAGGCACTCGAGGTGGTTCATCAGTAATATCTTCAACACAGACTACCATCAT 500  
 Db 519 GACCTCTCTCGTGTGAATGAGTTCATCTCTTAACTGCTTAAAGGTGAAGGATGCAGG 578

QY 501 GAACATGATGCACATCTACGTGTTCGACGCTATTTTGGCTGTCTGTGTGCTGGTGCCT 560  
 Db 579 AGGCTCCATGACCATCCACATTTTGGCGCTACTTTTGGCTCACAGTGAACCGGATCCT 638

QY 561 GCCAAGCCTTACCCGAGGGAACGGAGGATAAAGATCAGACAGCAACATACCCAGTTT 620  
 Db 639 CTACCCAGCAACCTAGACGAGCAAGGAGAGACAGAAATCTGTGTACCCAGTCGGACCT 698

QY 621 GTCTGCCATGCTGGGCGCCCTCTTCTGTGGATGTTCTGGCCCAAGTTTCAACTCTGCTCT 680  
 Db 699 CTTTGGCATGATGGCACCCCTCTTCTGTGGATGACTGGCCGAGCTTCAACTCAGCCAT 758

QY 681 GCTGAGAACTCCAATCGAAAGGAAGAATGCCGTGTTTCAACACTACTATGCTGTAGCACT 740  
 Db 759 ATCTTACCATGGGACAGCCAGCAGCCGAGCGCCCATCAACACTACTGCTCTTGGCAGC 818

QY 741 CAGCGTGTGACAGCATCTCAGGGTCACTTGGCTCACTCCGCCAAGGAAATCAGCAA 800  
 Db 819 CTGCGTGTCTTACCTCGTGGCAATATCCAGTCCCTCGACAAAGAGGGCAAGCTGGACAT 878

QY 801 GACTTATGTGCACAGTGGGCTGTTCGAGAGGCGTGGCTGGTGGTACCTCTGTTGACCT 860  
 Db 879 GGTGCACATCCAGAAATGCCACCTCGAGAGGGGTGGCGGTGGGTACCGCTGTGAGAT 938

QY 861 GATCCCTTTCGCTGGCTTGGCATGCTGTGGCTCTTGTGGCTGGGCTGATCTCGTGG 920  
 Db 939 GATGCTCATGCTTACGGTGGCTCATCATCGCTTCTGCGGCATCATCTCCACCT 998

QY 921 GGGAGCCAACTACCTGCGGGGTGTGTAAACGAGTGTGGGAGTCCCCACAGTCCCAT 980  
 Db 999 GGGTTTGTATACCTGAGCCCATTCCTGGAGTCCCGGCTGCACATCCAGGACACATGTGG 1058

QY 981 CATGGCTTACAACTTCACTGCTGGCTGTGGCTGTGGAGAGATCATCATATGTG 1035  
 Db 1059 CATTAACAATCTGCATGCTTCTGGCATCATAGGCGGCTCGTGGGTGCTGTG 1113

RESULT 12  
 ABK12334  
 ID ABK12334 standard; cDNA; 1679 BP.  
 XX AC ABK12334;  
 XX DT 05-JUN-2002 (first entry)  
 XX DE cDNA encoding mouse nonerythroid Rh glycoprotein Rhbg.  
 KW Mouse; nonerythroid Rh glycoprotein; Rhbg; RhBG; antibody; immunogen;  
 KW Rh type B transporter gene; polytypic transporter-type protein;  
 KW ion transporter; chromosome 3; gene; ss.  
 OS Mus sp.  
 FH Key Location/Qualifiers  
 FT CDS 39..1531  
 FT /\*tag= a  
 FT /product= "Mouse Rhbg protein"  
 FT polyA\_signal 1634..1639  
 FT /\*tag= b  
 FT /standard\_name= "polyA signal"  
 FT /note= "Atypical polyadenylation site"  
 XX WO200216396-A2.  
 XX PN 28-FEB-2002.  
 PD 17-AUG-2001; 2001WO-US025881.  
 XX PF 21-AUG-2000; 2000US-0226767P.  
 XX PR (NYBL-) NEW YORK BLOOD CENT INC.  
 XX PA  
 XX

PI Huang C, Liu Z;  
XX WPI: 2002-280905/32.  
DR P-PSDB; AAU78092.  
XX  
PT Homologs of mouse or human nonerythroid Rh glycoproteins, Rbpg or Rbpg.  
PT respectively, useful in the production of antibodies which are useful for  
PT detecting Rbpg or Rbpg glycoproteins in a sample.  
XX  
PS Claim 1; Fig 1a; 59pp; English.  
XX  
CC The present invention relates to a new protein or peptide comprising an  
CC amino acid sequence having at least 60% identity to a sequence comprising  
CC 455 (mouse nonerythroid Rh glycoprotein homologue (Rbpg)) or 458 (human  
CC nonerythroid Rh glycoprotein homologue (Rbpg)) amino acids, fully defined  
CC in the specification. The antibody of the invention is useful for  
CC detecting an Rbpg or an Rbpg glycoprotein in a sample, by contacting the  
CC sample with antibody under conditions suitable for binding, assessing the  
CC specific binding to the antibody, and thus detecting the presence of an  
CC epitope of Rbpg or Rbpg in the sample. The nucleic acids of the invention  
CC are useful as probes for detecting transporter genes and particularly Rh  
CC type B transporter genes including e.g. NH 4<sup>+</sup> ion transporters. The  
CC present nucleic acid sequence is that of the mouse Rbpg gene located on  
CC chromosome 3. This sequence encodes the mouse Rbpg protein of the  
CC invention. Rbpg is a polytypic transporter-type protein  
XX  
SQ Sequence 1679 BP; 324 A; 491 C; 469 G; 395 T; 0 U; 0 Other;  
Query Match 12.7%; Score 159.4; DB 6; Length 1679;  
Best Local Similarity 48.3%; Pred. No. 3.3e-35;  
Matches 445; Conservative 0; Mismatches 476; Indels 0; Gaps 0;  
141 CTATCAAGTTGGCGAAGATGACCGGTGATGGCGGCATTCGGCTTCTCCTCCTC 200  
209 CTACCAAGCTTCCAGGATGACGCTGATGCTTCTCGTGGGCTTGGCTTCTCCTCCT 268  
201 GAGTTTCCGAGACAGCTGGAGCGTGGCTTCAACTCTTCAATGCTGGCGCTTGG 260  
269 CTTTCTACAGCGTACGGCTTACGAGTGGGCTTACCTTCTCTGCGCCAGCTCAC 328  
261 TGTGCGTGGCGCAATCTGCTGGACGGCTTCTGAGCGAGTTCCTTCTGGGAGGTGGT 320  
329 CTTGCGAGTGGCGCACACTGCTCCAGGCTTCTTCCAGCTGGCGCACATCCA 388  
321 CATCACACTGTTGATGATTCGGCTGGCGCACATGAGTCTTGTGCGTGTGATCTCAGT 380  
389 TGTGCGTGGAGAGTTGATCAACGCTGACTTCTGCGGGGAGCTGTGCTCATCTCTT 448  
381 GGATGCTGCTTGGGGAGGTCAACTTGGGCGAGTGTGCTGATGCTGTGGTGGAGGT 440  
449 CGGGGCTGTTCTGGGCAAGACTGGGCGAGCCAGCTGCTGCTAATGCTTCTACTGGAGC 508  
441 GACAGCTTTAGGCAACCTGAGGATGGTCAATCAGTAATATCTTCAACACAGACTACCAT 500  
509 AGTGTGTTAGGTCAACGAGTTTATATCTACTCAGTCTCTCTGGGGGTGAGAGATGCTGG 568  
501 GAACATGATGATCACTACGTTGTCGAGCGCTATTTGGGCTGTGTCGCTGCTGCT 560  
569 AGGTGCTCAACAATCACATTTGGGGGCTTCTTCTGCGGCTGCTTCTCCTCAGGCTCT 628  
561 GCCAAGGCTCTACCCGAGGAGCGGAGGATAAAGATCAGACAGCAACGATACCCAGTTT 620  
629 CTACAGATCCAGCTGGAGAGAGCAGCATCGCAGAGCTCTGTCTACAACTCTGACCT 688  
621 GTCGTCATGCTGGGCGCCCTCTTCTTGTGAGTGTTCGCAAGTTTCAACTCTGCTCT 680  
689 CTTTGGCATGATCGGAGCACTTCTCTGTGGGTTTCTTGGGCGAGCTTCAACTCGGCC 748  
691 GCTGAGAGTCCAATCGAAGAGAAATGCGGTGTTCAACACCTTACTTGTGTAGCAGT 740  
749 GACAGCGTGGGGAGTGGGAGCATCGGACCGTGTGCAACATATCTTACTTACTCAGCG 808  
741 CAGCGTGTGACAGCATCTCAGGGTCTATCTGCTTGGCTCACCCCAAGGAGATCAGCAA 800

Db 809 AAGCACCCTCAGTACTTTTCGCTTGTCAAGCCCTTGTCAAGTGGAGATGGACGACAT 868  
QY 801 GACTTATGTGACAGTGGCTTGTGGCAGGAGCGTGGCTGGGTACCTGCTGTCACT 860  
Db 869 GGTCCACGCTCCAGAACGACGACACTGGCTGGAGGTGTGTGGTGGGACATCAAGTCAGAT 928  
QY 861 GATCCCTTCTCGTGGCTTGCATGCTGTGGTCTTGTGGCTGGGCTGATCTCCGTCGG 920  
Db 929 GATGTGACACCTTTGGGGCGCTGGCAGCTGGCTTCTGGCTGGGACTGTCTCCACACT 988  
QY 921 GGGAGCCAGTACCTGCCCGGGGTGTGTAAACGAGTGTGGGGATTCGCCACAGCTCCAT 980  
Db 989 GGGGTATAAGTTCTTTAGCCCTATCTTGAATCCAGATTAAACTGCAAGACACATGTGG 1048  
QY 981 CATGGCTACACATTCAGCTTCTGGTCTGTGGAGAGATCATCTACATTTGTGTGCT 1040  
Db 1049 TGTTCACACCTCCATGGGATGCCAGGGTCTCTGGGGGCACTCTGGGAGTCTGTAGTGGC 1108  
QY 1041 GGTGCTTGTATACCTGCGGAGC 1061  
Db 1109 TGCACCTGGCCACCCACGAGC 1129  
RESULT 13  
AAH18728/c  
ID AAH18728 standard; cDNA; 3706 BP.  
XX  
AC AAH18728;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human cDNA sequence SEQ ID NO:19002.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-00116126.  
XX  
PR 29-JUL-1999; 99JP-00248036.  
PR 27-AUG-1999; 99JP-00300253.  
PR 11-JAN-2000; 2000JP-00118776.  
PR 02-MAY-2000; 2000JP-00183767.  
PR 09-JUN-2000; 2000JP-00241899.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
DR WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.  
XX  
PS Claim 8; SEQ ID NO 19002; 2537pp + Sequence Listing; English.  
XX  
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
CC complementary strand of a polynucleotide which comprises one of the 5602  
CC nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
 CC specification. The primer sets can be used in antisense therapy and in  
 CC gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893  
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
 CC oligonucleotides, all of which are used in the exemplification of the  
 CC present invention  
 XX  
 SQ Sequence 3706 BP; 883 A; 1050 C; 684 G; 1089 T; 0 U; 0 Other;

Query Match 12.2%; Score 152.4; DB 4; Length 3706;  
 Best Local Similarity 99.4%; Pred. No. 5.4e-33;  
 Matches 153; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 333 CAGTATTCGGCTGGCCACCACATGAGTCTTTGTCGGTGTGATCTCATGTCAGTGTCTCTTT 392  
 Db 2906 CAGTATTCGGCTGGCCACCACATGAGTCTTTGTCGGTGTGATCTCATGTCAGTGTCTCTTT 2847  
 QY 393 GGGGAAGGTCAACTTGGCGCAGTGTGGTGTGATGTGCTGGTGGAGTGCACAGCTTTAGG 452  
 Db 2846 GGGGAAGGTCAACTTGGCGCAGTGTGGTGTGATGTGCTGGTGGAGTGCACAGCTTTAGG 2787  
 QY 453 CAACTGTAGAGTGTTCATCAGTAATATCTTCAAC 486  
 Db 2786 CAACTGTAGAGTGTTCATCAGTAATATCTTCAAC 2753

## RESULT 14

ABK49216  
 ID ABK49216 standard; cDNA; 2098 BP.  
 XX  
 AC ABK49216;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE cDNA encoding mouse Rh type C gene (rhcg) protein.  
 XX  
 KW RHCG: mouse; non-erythroid Rh type C glycoprotein; chromosome 7; gene;  
 KW ss.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 123..1496  
 FT /\*tag= a  
 FT /product= "rhcg protein"  
 FT 1934..1939  
 FT /\*tag= b  
 XX  
 FN WO200220719-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 05-SEP-2001; 2001WO-US027503.  
 XX  
 FR 07-SEP-2000; 2000US-0230660P.  
 XX  
 PA (NYBL-) NEW YORK BLOOD CENT INC.  
 XX  
 PI Huang C, Liu Z;  
 XX  
 DR WPI; 2002-351774/38.  
 DR P-PSDB; AAU78998.  
 XX  
 PT Nucleic acid sequences encoding novel mammalian nonerythroid Rh type C  
 PT and glycoproteins which have a characteristic twelve transmembrane domain  
 PT structure.

XX

Claim 3; Fig 1; 53pp; English.

This invention relates to the nucleic acid and protein sequences of novel human and mouse non-erythroid Rh type C glycoprotein (RhCG). The RhCG protein and the mouse homologue (rhcg) have a characteristic 12 transmembrane domain structure and are expressed in kidneys and testis. The invention also comprises a method for antibody that specifically binds an epitope of the glycoprotein and a method for detecting the protein using this antibody. The antibodies of the invention may be used in Western blots, enzyme linked immunosorbent assays (ELISA) or immunohistochemical assays to identify the non-erythroid tissues, particularly kidney and testis, that express the RhCG or Rhcg glycoproteins. The methods are used for detecting an Rhcg or RhCG glycoprotein in a sample. The present sequence represents the cDNA encoding the mouse RhCG Rh type C glycoprotein (RhCG) protein sequence of the invention. The gene encoding this protein is located on murine chromosome 7

SQ Sequence 2098 BP; 458 A; 594 C; 559 G; 487 T; 0 U; 0 Other;

Query Match 12.0%; Score 150.2; DB 6; Length 2098;  
 Best Local Similarity 47.5%; Pred. No. 1.7e-32;  
 Matches 446; Conservative 0; Mismatches 493; Indels 0; Gaps 0;  
 QY 97 CACTATGACGCTTCCTTAGAGGATCAAAAGGGGCTCGTGGCATCTCTATCAAGTTGGCCAA 156  
 Db 258 CGCAAGAACATCTCCAGCGACGTTGAGAACGAGTTTACTATCGCTTACCGAGCTTCCAG 317  
 QY 157 GATCTGACCGTATGGCGCCATTGGCTTGGGCTTCTCCTCAGCTCGAGTTTCGGAGACAC 216  
 Db 318 GATGTACACGCCATGGTCTTCTGCGGCTTCGCGCTTCTCTCATGACCTTCTCGAGCGCTAC 377  
 QY 217 AGCTGGAGCAGTGTGGCGCTTCAACCTCTTCATGCTGGCGCTTGGTGTGAGTGGGCAATC 276  
 Db 378 GGCTTCAGCGCTGTAGGCTTCACTTCTGCTGGCAGCTTTCGGCATCAGTGGGCACTG 437  
 QY 277 CTGCTGGACGGCTTCTCGAGCCAGTTCCCTTCTGGGAAGGTGGTCAATCACTGTTCACT 336  
 Db 438 CTCATGCGAGGATGGTTCCATTACATTGAAGAACCCACATTTGCTGAGCGTCGAGAAC 497  
 QY 337 ATTTCGGCTGGCCACCATGAGTGTCTTGTTCGGTGTGATCTCAGTGGATGCTGTCTTGGG 396  
 Db 498 ATCATCCAACTGACTTCTGTGTGGCATCTTCTGTTGGCTTCGCGCTTCGGGCGACTTCTAG 557  
 QY 397 AAGGTCAAACCTTGGCGCAGTGTGGTGTGATGTGTGGTGGAGTGCACAGCTTTAGGCAAC 456  
 Db 558 AAGGTCAAGCGGATGCAGCTGCTCATATGACCTTCTTCCAGTGAATCTCTTTCACAGTG 617  
 QY 457 CTGAGGATGGTCAATCAGTAATATCTTCAACACAGACTACCATGAAATGATGACATC 516  
 Db 618 AATGAGTTTCATCTCTGAACCTGATAGAGGCAAGGATGAGGGGGCTCTATGACCATC 677  
 QY 517 TAGCTGTTTCGAGCCTATTTTGGGCTGTCTGTGGCTGTGTGCTGCTGCCAAAGCTCTACCC 576  
 Db 678 CACACATTTGGCGCTACTTTGGGCTCACAGTGAACCTGGATCCTCTTACCGAAAAAACC 737  
 QY 577 GAGGGAACGAGGATAAAGATCAGACAGCAACGATACCCAGTGTGTCTGCCATGCTGGGC 636  
 Db 738 GATCAGAGCAAGCAGACAGACAGCTCAGTGTACCACCTCGGACCTTTTCGCCATGANTGGC 797  
 QY 637 GCCCTCTTCTTGTGGATGTTCTTGGCCAAAGTTTCAACTCTGCTGTGAGAAGTCCAATC 696  
 Db 798 ACCCTCTTCTTGTGGATATCTGGCCCAAGTTTCAATTTCAGCCAGTTCCTTCCACGGAGAT 857  
 QY 697 GAAAGGAGATGCGGTGTTCAACACCTACTATCTGTAGCAGTCAGGCTGGTGCAGGCC 756  
 Db 858 GCCCAGCACCGAGCGCCCTCAATACCTCTCTCTTGGCAGCGAGTGTCTTAACCAACA 917  
 QY 757 ATCTCAGGCTCATCTTGGCTCACCCCAAGGGAGATCAGCAAGACTTATGTGCACAGT 816  
 Db 918 GTGACAGTATCCAGTATTTGTACACAAGAGGGCAAGTTGGATATGGTGACATCCAGAT 977





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 12:40:53 ; Search time 3117 Seconds

(without alignments)  
12013.851 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254

Sequence: 1 atgagctctaagtaaccgcg.....atttgctgttgattttaa 1254

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_nam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gsl:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	824.4	65.7	994	13	BX406635
2	633.8	50.5	685	13	BX406635
3	629.8	50.2	676	13	BX106517
4	617.2	49.2	685	13	BX406635

C	5	609	48.6	1000	13	BX406634
6	602.2	48.0	1457	11	AK079335	Mus muscu
7	597.4	47.6	706	13	BU661907	cl178e10.z
8	560.4	44.7	631	13	BU662207	cl182a04.z
9	559	44.6	632	13	BU665450	cl131a04.z
10	552	43.3	607	13	BU664659	cl120e10.z
11	542.4	39.8	607	13	BU657533	Mus muscu
12	498.6	39.8	975	11	AK089642	cl125d06.z
13	427.2	34.1	612	13	BU663660	cl104e12.z
14	405.6	32.3	587	13	BU661598	cl174e01.z
15	399.2	31.8	480	13	BU659596	cl149f09.z
16	398.6	31.8	683	12	BM719724	cl149f09.z
17	392.4	31.3	443	9	AL699045	cl149f09.z
18	359	28.6	429	13	BU657883	cl129d10.z
19	349.2	27.8	721	13	BY750607	cl129d10.z
20	343.2	27.4	551	12	BM254191	cl129d10.z
21	321.2	25.6	660	13	BY722497	cl129d10.z
22	317.6	25.3	643	12	BM684087	cl129d10.z
23	314.4	25.1	701	13	BY734317	cl129d10.z
24	302.2	24.1	434	14	N59044	cl129d10.z
25	299.6	23.9	375	13	BU664580	cl129d10.z
26	294.2	23.5	499	10	BF603905	cl129d10.z
27	288.2	23.0	488	10	BF602079	cl129d10.z
28	280.4	22.4	560	29	CG573338	cl129d10.z
29	273.6	21.8	515	13	EX529358	cl129d10.z
30	271.2	21.6	498	12	BG944752	cl129d10.z
31	270.4	21.6	416	14	R10548	cl129d10.z
32	257.8	20.6	601	13	BU661774	cl129d10.z
33	255.8	20.4	522	29	CG634208	cl129d10.z
34	253.6	20.2	343	14	T84327	cl129d10.z
35	250	19.9	504	29	CG669296	cl129d10.z
36	243	19.4	515	10	BF191605	cl129d10.z
37	240.8	19.2	474	13	BY563146	cl129d10.z
38	240.8	19.2	476	13	BY559962	cl129d10.z
39	240.8	19.2	476	13	BY592375	cl129d10.z
40	240.2	19.2	522	29	CG570751	cl129d10.z
41	239.8	19.1	501	29	CG558871	cl129d10.z
42	236.8	18.9	491	9	AA259801	cl129d10.z
43	231.8	18.5	446	10	BB847808	cl129d10.z
44	230.2	18.4	461	13	BY557825	cl129d10.z
45	225.8	18.0	453	13	BY591481	cl129d10.z

## ALIGNMENTS

RESULT 1	BX406635	934 bp	linear	EST 15-MAY-2003
LOCUS	BX406635	Homo sapiens FETAL LIVER	Homo sapiens	cdna clone
DEFINITION	CSODM013YA17 5-PRIME, mRNA sequence.			
ACCESSION	BX406635			
VERSION	BX406635.1	GI:30762708		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Li W.B., Gruber C., Jessee J., and Polayes D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 854.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CSODM013AA09QP1&cluster=854.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600			

Paraday Avenue Genoscope sequence ID : CS0AM013AA09P1.

FEATURES

Location/Qualifiers  
1. .994  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DM013YA17"  
/tissue\_type="FETAL LIVER"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL LIVER"  
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."  
Query Match 65.7%; Score 824.4; DB 13; Length 994;  
Best Local Similarity 93.0%; Pred. No. 7.3e-198;  
Matches 857; Conservative 16; Mismatches 47; Indels 2; Gaps 1;  
QY 1 ATGAGCTCTAAGTACCGCGGTCTGTCGGCGTCTGTCGGCGTCTGCGCCCTTAACACTG 60  
Db 73 ATGAGCTCTAAGTACCGCGGTCTGTCGGCGTCTGTCGGCGTCTGCGCCCTTAACACTG 132  
QY 61 GAAGCAGCTCTCATCTCCCTCTTCTATTTTACCCACTATGACGCTTCTCTTAGAGGAT 120  
Db 133 GAAGCAGCTCTCATCTCCCTCTTCTATTTTBBBACCCTATGACGCTTCTCTTAGAGGAT 192  
QY 121 CAAAGGGGCTCGGCGATCTATCAAGTTGGCCAGATCTGACCGTGTGCGGCCCAT 180  
Db 193 CAAAGGGGCTCGGCGATCTATCAAGTTGGCCAGATCTGACCGTGTGCGGCCCAT 252  
QY 181 GGCTTGGGCTTCTCACTCGAGTTTCGGAGACACAGCTGGAGAGTGGCCCTTCAAC 240  
Db 253 GGCTTGGGCTTCTCACTCGAGTTTCGGAGACACAGCTGGAGAGTGGCCCTTCAAC 312  
QY 241 CTCTTCATGCGGCTGGTGTGTCAGTGGGCAATCTGTCGGACGCTTCTCTGAGCAG 300  
Db 313 CTCTTCATGCGGCTGGTGTGTCAGTGGGCAATCTGTCGGACGCTTCTCTGAGCAG 372  
QY 301 TTCCCTTCTGGAGGTTGGTCTATCACACTGTTCAGTATTCGGCTGGCCACCATGAGTGT 360  
Db 373 TTCCCTTCTGGAGGTTGGTCTATCACACTGTTCAGTATTCGGCTGGCCACCATGAGTGT 432  
QY 361 TTGTCTGTGCTGATCTCAGTGGAGTCTCTTGTGGGAGGTCACCTGGCCGAGTTGGT 420  
Db 433 ATGTCTGGTGTGATCTCAGCGGGTCTCTTGTGGGAGGTCACCTGGCCGAGTTGGT 492  
QY 421 GTGATGTGTGGTGGAGTGACAGCTTTAGGCACTGAGGATGGTCTATCAGTAATATC 480  
Db 493 GTGATGTGTGGTGGAGTGACAGCTTTAGGCACTGAGGATGGTCTATCAGTAATATC 552  
QY 481 TTCAACACAGACTACCATGAACATGATGACATCTACGTTTTCGACGCTATTTTGG 540  
Db 553 TTCAACACAGACTACCATGAACATGATGACATCTACGTTTTCGACGCTATTTTGG 612  
QY 541 CTGTCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 613 CTGTCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672  
QY 601 ACAGCAACGATACCCAGTTTGTCTCCATGCTGGGCGCCCTCTTCTGTGGATGTTCTGG 660  
Db 673 AGAGCAACGATACCCAGTTTGTCTCCATGCTGGGCGCCCTCTTCTGTGGATGTTCTGG 732  
QY 661 CCAAGTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720  
Db 733 CCAAGTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 792  
QY 721 ACCTACTATGCTGTAGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 780  
Db 793 ACCTACTATGCTGTAGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 852

ORIGIN

BU664461 685 bp mRNA linear EST 30-SEP-2002  
Homo sapiens cDNA clone c1117b02 5', mRNA sequence.  
BU664461  
EST.  
BU664461.1 GI:23376646

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 685)  
Gubin, A.N., Lee, Y.T., Bouffard, G.G. and Miller, J.L.  
Gene Expression in Human Erythroid Precursor Cells  
Unpublished (2002)  
Contact: Jeffery L. Miller  
Laboratory of Chemical Biology  
National Institute of Diabetes and Digestive and Kidney Diseases  
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD 20892, USA  
Tel: 301 402 2373  
Fax: 301 435 5148  
Email: jmf@nih.gov  
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or analyses by National Institutes of Health Intramural Sequencing Center (NISC). More information available at:  
http://hembase.nidk.nih.gov  
Plate: 117 row: b column: 02  
Seq primer: 5' lambda-TripleX2 Sequencing Primer.  
Location/Qualifiers

FEATURES

1. .685  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="c1117b02"  
/sex="unknown"  
/tissue\_type="blood"  
/cell\_type="Erythroid Precursor Cells"  
/cell\_line="Primary Culture of Peripheral Blood Mononuclear Cells"  
/dev\_stage="Precursor erythroblasts; GPA++"  
/lab\_host="DH5alpha"  
/clone\_lib="Hembase; Erythroid Precursor Cells (LCB:cl library)"  
/note="Organ: blood; Vector: pTriplex2; Site 1: Sfil;  
Site 2: Sfil; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA Library Construction Kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos  
(N-terminal)-biotin-GTC-CAC-CCG-AGG-CTT-G-(C-terminal) and  
(N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal)  
. Synthesized cDNA was digested with Sfil and  
size-selected on a 1% agarose gel (>800bp). Large-scale  
sequencing of the library was performed by the NIH

QY 781 CCCAAGGAGAGATCAGCAAGACTTATGTGCACAGTGCCTGTGTGCAGAGCGGTGGCT 840  
Db 853 CCCAAGGAGAGATCAGCAAGACTTATGTGCACAGTGCCTGTGTGCAGAGCGGTGGCT 912  
QY 841 GTGGGTACCTCG--TGTACCTGATCCCTTCTCCGTGGCTTGCATGCTCGGTCTTG 898  
Db 913 GTGGGTACTCGTGTCTACCTGAKCCCTTCTCGGGGGCTTGCASWSGTGCTCGGTCTC 972  
QY 899 TGGCTGGCTGATCTCCGTGG 920  
Db 973 TTGGCTGGGTGATTGCAHCGG 994





```

http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOAM013AA09NP1&cluster=854.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSOAM013AA09NP1.
Location/Qualifiers
1. 1000
/organism="Homo sapiens"
/mol_type="tRNA"
/db_xref="taxon:9606"
/clone="CSODM013YA17"
/tissue_type="FETAL LIVER"
/dev stage="fetal"
/clone.lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

[illegible]



```

DEFINITION cl78e10.z1 Hembase; Erythroid Precursor Cells (ICB:cl library) Homo
sapiens cDNA clone cl78e10 5', mRNA sequence.
ACCESSION BU661907
VERSION BU661907.1 GI:23374089
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 706)
Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
Gene Expression in Human Erythroid Precursor Cells
Unpublished (2002)
Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 402 2373
Fax: 301 435 5148
Email: jm7f@nih.gov
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing
Center (NISC). More information available at:
http://hembase.niddk.nih.gov
Plate: 78 row: e column: 10
Seq primer: 5' lambda-TripleX2 Sequencing Primer.
Location/Qualifiers
1. 706
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cl78e10"
/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood
Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GPA++"
/lab_host="DH5alpha"
/clone_lib="Hembase; Erythroid Precursor Cells (ICB:cl
library)"
/notes="Organ: blood; Vector: pTriplex2; Site 1: SfiI;
Site 2: SfiI; A complementary DNA (cDNA) library from
human erythroid precursor cells was constructed using
SMART PCR (polymerase chain reaction) cDNA Library
Construction Kit (Clontech, Palo Alto, CA) according to
the manufacturer's directions, but with slight
modifications. Briefly, reverse transcription was
performed in the presence of 1 umol/L peptide nucleic acid
(PNA) oligos
(N-terminal)-biotin-GTC-CAC-CCG-AAG-CTT-G-(C-terminal) and
(N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal)
. Synthesized cDNA was digested with SfiI and
size-selected on a 1% agarose gel (>800bp). Large-scale
sequencing of the library was performed by the NIH
Intramural Sequencing Center (NISC;
http://www.nisc.nih.gov/)."
FEATURES
source

```

## ORIGIN

Query Match 47.6%; Score 597.4; DB 13; Length 706;  
Best Local Similarity 97.4%; Pred. No. 2.5e-140;  
Matches 607; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

1	ATGAGCTCTAAGTACCCGGGTCTGTGCGCGCTGCTGCGCCCTCTGGGGCCCTAAACACTG	60
		Db
84	ATGAGCTCTAAGTACCCGGGTCTGTGCGCGCTGCTGCGCCCTCTGGGGCCCTAAACACTG	143
		Qy
61	GAAGCAGCTCTCAATTCCTCTCTATTTTTTTTACCACCTATGACGCTTCCTTTAGAGAT	120
		Db
144	GAAGCAGCTCTCAATTCCTCTCTATTTTTTTTACCACCTATGACGCTTCCTTTAGAGAT	203

QY	121	CAAAAGGGGCTCGTGGCATCCTATCAAGTTGGCCACAGATCTGACCGTGATGGCGGCCATT	180
Db	204	CAAAAGGGGCTCGTGGCATCCTATCAAGTTGGCCACAGATCTGACCGTGATGGCGGCCCTT	263
QY	181	GGCTTGGGCTTCCTCACTCGAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCCCTTCAAC	240
Db	264	GGCTTGGGCTTCCTCACTCAAAATTTCCGGAGACACAGCTGGAGCAGTGTGGCCCTTCAAC	323
QY	241	CTCTTCATGCTGGCGCTTGGTGTGCGATGGGCAATCTGTGGAACGCTTCTTGAGCCAG	300
Db	324	CTCTTCATGCTGGCGCTTGGTGTGCGATGGGCAATCTGTGGAACGCTTCTTGAGCCAG	383
QY	301	TTCCCTTCTGGGAAGTGGTCATCACACTGTTTCAGTATTCCGCTGGCCACCATGAGTGCT	360
Db	384	TTCCCTTCTGGGAAGTGGTCATCACACTGTTTCAGTATTCCGCTGGCCACCATGAGTGCT	443
QY	361	TTTTCGGTGTGTGATCTCAGTGGATGCTGTCTTGGGAAAGTCAACTTGGCGCAGTTGGTG	420
Db	444	ATGTCGGTGTGTGATCTCAGCGGTGCTGTCTTGGGAAAGTCAACTTGGCGCAGTTGGTG	503
QY	421	GTGATGGTGTGGTGGAGGTGACAGCTTTTAGGCAACCTGAGGATGGTCATCAGTAATATC	480
Db	504	GTGATGGTGTGGTGGAGGTGACAGCTTTTAGGCAACCTGAGGATGGTCATCAGTAATATC	563
QY	481	TTCAACACAGACTACCAATGACATGATGCAATCTACGTGTTTCGAGCCTATTTTGGG	540
Db	564	TTCAACACAGACTACCAATGACATGATGCAATCTACGTGTTTCGAGCCTATTTTGGG	623
QY	541	CTCTCTGTGCGCTGTGCTGCCAAAGCCTCTACCCGAGGGAACGGAGGATAAAGATCAG	600
Db	624	CTGACTGTGCGCTGTGCTGCCAAAGCCTCTACCCGAGGGAACGGAGGATAAAGATCAG	683
QY	601	ACAGCAACGATACCCAGTTTGTC	623
Db	684	AGAGCAACGATACCCAGTTTGTC	706
RESULT	8		
BU652207			
LOCUS		631 bp	linear
DEFINITION		EST 30-SEP-2002	
		cl182h04.z1 Hembase; Brythroid Precursor Cells (iCB:cl library) Homo	
		sapiens cDNA clone cl182h04.5', mRNA sequence.	

LOCUS	BU662207	631 bp	linear	EST 30-SEP-2002
DEFINITION	cl82h04.z1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo sapiens cDNA clone cl82h04 5', mRNA sequence.			

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	







BU657533  
LOCUS 607 bp mRNA linear EST 30-SEP-2002  
DEFINITION sapiens cdNA clone c125d06 5', mRNA sequence.  
ACCESSION BU657533  
VERSION BU657533.1 GI:23369715  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 607)  
AUTHORS Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.  
TITLE Gene Expression in Human Erythroid Precursor Cells  
JOURNAL Unpublished (2002)  
COMMENT Contact: Jeffery L. Miller  
Laboratory of Chemical Biology  
National Institute of Diabetes and Digestive and Kidney Diseases  
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD  
20892, USA  
Tel: 301 402 2373  
Fax: 301 435 5148  
Email: jmf@nih.gov  
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the  
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or  
analyses by National Institutes of Health Intramural Sequencing  
Center (NISC). More information available at:  
http://hembase.nidk.nih.gov  
Plate: 25 row: d column: 06  
Seq primer: 5' lambda-Triplex2 Sequencing Primer.  
FEATURES  
Location/Qualifiers  
1..607  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="c125d06"  
/sex="unknown"  
/tissue\_type="blood"  
/cell\_type="Erythroid Precursor Cells"  
/cell\_line="Primary Culture of Peripheral Blood  
Mononuclear Cells"  
/dev\_stage="Precursor erythroblasts; GPA++"  
/lab\_host="DH5alpha"  
/clone\_lib="Hembase; Erythroid Precursor Cells (LCB:cl  
library)"  
/note="Organ: blood; Vector: pTriplex2; Site: 1: Sfil;  
Site: 2: Sfil; A complementary DNA (cDNA) library from  
human erythroid precursor cells was constructed using  
SMART PCR (polymerase chain reaction) cDNA Library  
Construction Kit (Clontech, Palo Alto, CA) according to  
the manufacturer's directions, but with slight  
modifications. Briefly, reverse transcription was  
performed in the presence of 1 umol/L peptide nucleic acid  
(PNA) oligos  
(N-terminal)-biotin-GTC-CAC-AGG-CTT-G-(C-terminal) and  
(N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal)  
. Synthesized cDNA was digested with Sfil and  
size-selected on a 1% agarose gel (>800bp). Large-scale  
sequencing of the library was performed by the NIH  
Intramural Sequencing Center (NISC;  
Http://www.nisc.nih.gov/)."

Query Match 43.3%; Score 542.4; DB 13; Length 607;  
Best Local Similarity 98.0%; Pred. No. 2.1e-126;  
Matches 549; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
Qy 1 ATGAGCTCTAAGTACCGCGTCTGTCCGCGTGTCTGCGCCCTCTGCGCCCTAAACACTG 60  
Db 48 ATGAGCTCTAAGTACCGCGTGTCTGTCCGCGTGTCTGCGCCCTCTGCGCCCTAAACACTG 107  
Qy 61 GAGCAGCTTCATCTCCCTCTCTATTTTACCACATGATGACGCTTCCTTAGAGAT 120  
|||||

Db 108 GAGCAGCTTCATCTCCCTCTCTATTTTACCACATGATGACGCTTCCTTAGAGAT 167  
Qy 121 CAAAGGGGCTCTGTCATCTCTCAAGTTGGGCAAGATCTGACCGTGATGGGGCCATT 180  
|||||  
Db 168 CAAAGGGGCTCTGTCATCTCTCAAGTTGGGCAAGATCTGACCGTGATGGGGCCATT 227  
|||||  
Qy 181 GGCTTGGGCTTCCTCACCTCGAGTTTCCGGAGACACAGCTGGAGCAGTGGGCTTCAAC 240  
Db 228 GGCTTGGGCTTCCTCACCTCGAGTTTCCGGAGACACAGCTGGAGCAGTGGGCTTCAAC 287  
Qy 241 CTCCTTCATGCTGGCGCTTGGTGTGCACTGGGCAATCTCTGTCGACCGCTTCCTGAGCAG 300  
Db 288 CTCCTTCATGCTGGCGCTTGGTGTGCACTGGGCAATCTCTGTCGACCGCTTCCTGAGCAG 347  
Qy 301 TTCCCTTCTGGGAAGGTGGTGCATCACACTGTTTCACTATTCGGCTGGCCACCATGAGTCT 360  
Db 348 TTCCCTTCTGGGAAGGTGGTGCATCACACTGTTTCACTATTCGGCTGGCCACCATGAGTCT 407  
Qy 361 TTCTCGGTGCTGATCTTCAGTGGATGCTGCTTGGGGAAGGTCAACTTGGGCGAGTTGGTG 420  
Db 408 ATGTCGGTGTGATCTTCAGGCGGTGCTGCTTGGGGAAGGTCAACTTGGGCGAGTTGGTG 467  
Qy 421 GTCATGCTGCTGTGAGGTGACAGCTTTAGGCAACCTGAGGATGTCATCAGTAATATC 480  
Db 468 GTGATGCTGCTGTGAGGTGACAGCTTTAGGCAACCTGAGGATGTCATCAGTAATATC 527  
Qy 481 TTCAACACAGACTTACCACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
Db 528 TTCAACACAGACTTACCACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 587  
Qy 541 CTGCTGTGTGGCGCTGGTGCCT 560  
Db 588 CTGCTGTGTGGCGCTGGTGCCT 607  
AK089642 975 bp mRNA linear HTC 20-SEP-2003  
Mus musculus activated spleen cDNA, RIKEN full-length enriched  
library, clone:F83006J05 product:Rhesus blood group CE and D, full  
insert sequence.  
LOCUS AK089642  
DEFINITION AK089642.1 GI:26354620  
VERSION AK089642.1  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
AUTHORS Carninci,P. and Hayashizaki,Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE 2  
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE 3  
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
Konno,H., Akiyama,J., Nishi,K., Kutsunai,T., Tashiro,H., Itoh,M.,  
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,  
Yamanoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)



```

RESULT 13
BU663660
LOCUS
DEFINITION
  BU663660 612 bp mRNA linear EST 30-SEP-2002
  cl104e12.z1 Hembase; Erythroid Precursor Cells (LCB:cl library)
ACCESSION
  BU663660
VERSION
  BU663660.1 GI:23375845
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 612)
AUTHORS
  Gubin,A.N., Lee,Y.T., Bouffard G.G. and Miller,J.L.
TITLE
  Gene Expression in Human Erythroid Precursor Cells
JOURNAL
  Unpublished (2002)
COMMENT
  Contact: Jeffery L. Miller
  Laboratory of Chemical Biology
  National Institute of Diabetes and Digestive and Kidney Diseases
  Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
  20892, USA
  Tel: 301 402 2373
  Fax: 301 435 5148
  Email: jm7@nih.gov
  The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
  Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
  analyses by National Institutes of Health Intramural Sequencing
  Center (NISC). More information available at:
  http://hembase.nidk.nih.gov
  Plate: 104 row: e column: 12
  Seq primer: 5' lambda-Triplex2 Sequencing Primer.
FEATURES
  Location/Qualifiers
    1..612
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="cl104e12"
      /sex="unknown"
      /tissue_type="blood"
      /cell_type="Erythroid Precursor Cells"
      /cell_line="Primary Culture of Peripheral Blood
      Mononuclear Cells"
      /dev_stage="Precursor erythroblasts; GPA++"
      /lab_host="DH5alpha"
      /clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl
      library)"
      /notes="Organ: blood; Vector: pTriplex2; Site_1: Sfil;
      Site_2: Sfil; A complementary DNA (cDNA) library from
      human erythroid precursor cells was constructed using
      SMART PCR (polymerase chain reaction) cDNA Library
      Construction Kit (Clontech, Palo Alto, CA) according to
      the manufacturer's directions, but with slight
      modifications. Briefly, reverse transcription was
      performed in the presence of 1 umol/L peptide nucleic acid
      (PNA) oligos
      (N-terminal)-biotin-GTC-CAC-CCG-AAG-CTT-G-(C-terminal) and
      (N-terminal)-biotin-C(T/C)T-GAA-CTT-CTC-AGG-A-(C-terminal)
      . Synthesized cDNA was digested with Sfil and
      size-selected on a 1% agarose gel (>800bp). Large-scale
      sequencing of the library was performed by the NIH
      Intramural Sequencing Center (NISC;
      Http://www.nisc.nih.gov/)."
ORIGIN
  Query Match 34.1%; Score 427.2; DB 13; Length 612;
  Best Local Similarity 89.6%; Pred. No. 3.6e-97;
  Matches 489; Conservative 0; Mismatches 13; Indels 44; Gaps 1;
  QY 6 CTCCTAAGTACCCGGGTCTGTCGGGGCTGCTGCTGGCCCTCTGGGCCCTACACTGGAAGC 65
  Db 67 CTCCTAAGTACCCGGGTCTGTCGGGGCTGCTGCTGGCCCTCTGGGCCCTACACTGGAAGC 126
  QY 66 AGCTCTCATTCTCTCTTATTATTTTATACCACTATGAGCTTCCTTAGAGGATCAAAA 125

```

```

Db 127 AGCTCTCATTCTCTCTTATTATTTTATACCACTATGAGCTTCCTTAGAGGATCAAAA 186
QY 126 GGGGCTCGTGGGATCCTATC----- 145
Db 187 GGGGCTCGTGGGATCCTATC----- 246
QY 146 ----AAAGTTGGCCAAAGATCTGACCGTGGGGCCATTGGCTGGGCTTCCTCAGCTCG 201
Db 247 CCAGAAAGTGGCCAAAGATCTGACCGTGGGGCCCTTGGCTGGGCTTCCTCAGCTCA 306
QY 202 AGTTTCGGGAGACACAGCTGGAGCAGTGTGGCCTTCAACCTCTTTCATGCTGGCGCTGGT 261
Db 307 AATTTCGGGAGACACAGCTGGAGCAGTGTGGCCTTCAACCTCTTTCATGCTGGCGCTGGT 366
QY 262 GTGCAGTGGGCAATCTCTGTGGACGGCTTCTGAGCCAGTTCCCTCTGCGAAGGTGGTC 321
Db 367 GTGCAGTGGGCAATCTCTGTGGACGGCTTCTGAGCCAGTTCCCTCTGCGAAGGTGGTC 426
QY 322 ATCACACTGTTCAGTATTCGGCTGGCCACCATGAGTGTTCGGTGTGCTGATCTCAGTG 381
Db 427 ATCACACTGTTCAGTATTCGGCTGGCCACCATGAGTGTTCGGTGTGCTGATCTCAGTG 486
QY 382 GATGCTGTCTTGGGGAAGGTCAACTTGGGCGCAGTTGGTGGTGTGCTGCTGGTGGAGTG 441
Db 487 GGTGCTGTCTTGGGGAAGGTCAACTTGGGCGCAGTTGGTGGTGTGCTGCTGGTGGAGTG 546
QY 442 ACAGCTTTAGGCACCTGAGGATGTCATCAGTAATATCTTCAACACAGACTACACATG 501
Db 547 ACAGCTTTATGCACCTGAGGATGTCATCAGTAATATCTTCAACACAGACTACACATG 606
QY 502 AACATG 507
Db 607 AACCTG 612

```

```

RESULT 14
BU661598
LOCUS
DEFINITION
  BU661598 587 bp mRNA linear EST 30-SEP-2002
  cl74c01.z1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo
  sapiens cDNA clone cl74c01 5', mRNA sequence.
ACCESSION
  BU661598
VERSION
  BU661598.1 GI:23373780
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 587)
AUTHORS
  Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
TITLE
  Gene Expression in Human Erythroid Precursor Cells
JOURNAL
  Unpublished (2002)
COMMENT
  Contact: Jeffery L. Miller
  Laboratory of Chemical Biology
  National Institute of Diabetes and Digestive and Kidney Diseases
  Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
  20892, USA
  Tel: 301 402 2373
  Fax: 301 435 5148
  Email: jm7@nih.gov
  The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
  Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
  analyses by National Institutes of Health Intramural Sequencing
  Center (NISC). More information available at:
  http://hembase.nidk.nih.gov
  Plate: 74 row: c column: 01
  Seq primer: 5' lambda-Triplex2 Sequencing Primer.
FEATURES
  Location/Qualifiers
    1..587
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="cl74c01"

```

/sex="unknown"  
/issue\_type="blood"  
/cell\_type="Erythroid Precursor Cells"  
/cell\_line="Primary Culture of Peripheral Blood Mononuclear Cells"  
/dev\_stage="Precursor erythroblasts; GPA++"  
/lab\_host="DH5alpha"  
/clone\_lib="Hembase; Erythroid Precursor Cells (LCB:cl library)"  
/note="Organ: blood; Vector: pTriplex2; Site.1: SfiI; Site.2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA Library Construction Kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos  
(N-terminal)-biotin-GTC-CAC-CCG-AAG-CTT-G-(C-terminal) and  
(N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal). Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC;  
Http://www.nisc.nih.gov/).

## ORIGIN

Query Match 32.3%; Score 405.6; DB 13; Length 587;  
Best Local Similarity 89.8%; Pred. No. 1.1e-91;  
Matches 465; Conservative 0; Mismatches 9; Indels 44; Gaps 1;  
QY 1 ATGAGCTCTAAGTACCGCGGCTGTGCGCGCTGCTGCCCTCTGGCCCTTAACACTG 60  
Db |||||  
QY 70 ATGAGCTCTAAGTACCGCGGCTGTGCGCGCTGCTGCCCTCTGGCCCTTAACACTG 129  
Db |||||  
QY 61 GAAGCAGCTCTCATTTCTCTCTTCTATTTTACCACCTATGACGCTTCTTAGAGAT 120  
Db |||||  
QY 130 GAAGCAGCTCTCATTTCTCTCTTCTATTTTACCACCTATGACGCTTCTTAGAGAT 189  
QY 121 CAAAGGGGCTGTGGCATCTCTATC----- 145  
Db |||||  
QY 190 CAAAGGGGCTGTGGCATCTCTATCAAGGGGCTGGGAAGTCTGCATGCTCTATAAATC 249  
QY 146 -----AGTTGCCAGATCTCACCTGATGGGCGCATGCTGGGCTTCTCA 196  
Db |||||  
QY 250 CAGAACCAAGTCTGGCCAGATCTCACCTGATGGGCGGCTTGGGCTTCTCA 309  
QY 197 CTTGAGTTTCCGAGACACAGCTGGAGCAGTGTGGCCTTCAACCTCTTTCATGCTGGCG 256  
Db |||||  
QY 310 CTTCAATTTCCGAGACACAGCTGGAGCAGTGTGGCCTTCAACCTCTTTCATGCTGGCG 369  
QY 257 TTGGTGTGAGTGGGCAATCTGTCGACGCTTCTGACGCACTGTCCTTCTGGGAAG 316  
Db |||||  
QY 370 TTGGTGTGAGTGGGCAATCTGTCGACGCTTCTGACGCACTGTCCTTCTGGGAAG 429  
QY 317 TGGTCAACACTGTTTCAATTCGCTGGCCACCATGATGCTTGTGCGTCTGATCT 376  
Db |||||  
QY 430 TGGTCAACACTGTTTCAATTCGCTGGCCACCATGATGCTTGTGCGTCTGATCT 489  
QY 377 CAGTGGATGCTGTCTTGGGAAGTCAACTTGGCGCAGTTGTTGTTGATGTTGCTGGG 436  
Db |||||  
QY 490 CAGCGGTGCTGTCTTGGGAAGTCAACTTGGCGCAGTTGTTGTTGATGTTGCTGGG 549  
QY 437 AGTGTACAGCTTAGGCACCTGAGATGTTATCAGT 474  
Db |||||  
QY 550 AGTGTACAGCTTAGGCACCTGAGATGTTATCAGT 587

## RESULT 15

BUG59596

LOCUS

DEFINITION BU659596 480 bp mRNA linear EST 30-SEP-2002  
cl49f09.z1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo sapiens cDNA clone cl49f09 5', mRNA sequence.

ACCESSION BU659596  
VERSION BU659596.1 GI:23371778  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 480)  
Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.  
Gene Expression in Human Erythroid Precursor Cells  
Unpublished (2002)  
Contact: Jeffery L. Miller  
Laboratory of Chemical Biology  
National Institute of Diabetes and Digestive and Kidney Diseases  
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD 20892, USA  
Tel: 301 402 2373  
Fax: 301 435 5148  
Email: jml7@nih.gov  
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or analyses by National Institutes of Health Intramural Sequencing Center (NISC). More information available at:  
http://hembase.nidk.nih.gov  
Plate: 49 row: f column: 09  
Seq primer: 5' lambda-triPLEX2 Sequencing Primer.  
Location/Qualifiers  
1. .480  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="cl49f09"  
/sex="unknown"  
/tissue\_type="blood"  
/cell\_type="Erythroid Precursor Cells"  
/cell\_line="Primary Culture of Peripheral Blood Mononuclear Cells"  
/dev\_stage="Precursor erythroblasts; GPA++"  
/lab\_host="DH5alpha"  
/clone\_lib="Hembase; Erythroid Precursor Cells (LCB:cl library)"  
/note="Organ: blood; Vector: pTriplex2; Site.1: SfiI; Site.2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA Library Construction Kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos  
(N-terminal)-biotin-GTC-CAC-CCG-AAG-CTT-G-(C-terminal) and  
(N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal). Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC;  
Http://www.nisc.nih.gov/).

FEATURES  
source

## ORIGIN

Query Match 31.8%; Score 399.2; DB 13; Length 480;  
Best Local Similarity 98.1%; Pred. No. 4.1e-90;  
Matches 404; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 ATGAGCTCTAAGTACCGCGGCTGTGCGGCGCTGCTGCCCTCTGGCCCTTAACACTG 60  
Db |||||  
QY 69 ATGAGCTCTAAGTACCGCGGCTGTGCGGCGCTGCTGCCCTCTGGCCCTTAACACTG 128  
Db |||||  
QY 61 GAAGCAGCTCTCATTTCTCTCTTCTATTTTACCACCTATGACGCTTCTTAGAGAT 120  
Db |||||  
QY 129 GAAGCAGCTCTCATTTCTCTCTTCTATTTTACCACCTATGACGCTTCTTAGAGAT 188  
QY 121 CAAAGGGGCTGTGGCATCTCTATCAAGTTGGCCAGATCTGACGCTGATGGGCCATT 180  
Db |||||

```

Db      189  CAAAGGGGCTCGTGGCATCCTATCAAGTCGGCCAGATCTGACCGTGATGGCGGCCCTT 248
QY      181  GGCTTGGGCTTCCTCACCCTCGAGTTTCCGAGAGACACAGCTGGAGCAGTGTGGCCTTCAAC 240
Db      249  GGCTTGGGCTTCCTCACCCTCAAAATTCGGGAGACACAGCTGGAGCAGTGTGGCCTTCAAC 308
QY      241  CTCTTCATGCTGGGCTTGGTGTGCACTGGGCAATCCTGCTGGACGGCTTCTTGAGCCAG 300
Db      309  CTCTTCATGCTGGGCTTGGTGTGCACTGGGCAATCCTGCTGGACGGCTTCTTGAGCCAG 368
QY      301  TTCCCTTCTGGGAAGTGTGTCATCACACTGTTCAAGTATTCGGCTGGCCACCACATGAGTCT 360
Db      369  TTCCCTCCTGGGAAGTGTGTCATCACACTGTTCAAGTATTCGGCTGGCCACCACATGAGTCT 428
QY      361  TTGTCGGTGTGATCTCAGTGGATGCTGCTTGGGGAGGTCAACTTGGGCGC 412
Db      429  ATGTGGTGTGATCTCAGCGGGTGTGTCTTGGGGAAGGTCAACTTGGGCGC 480

```

Search completed: August 25, 2004, 17:28:29  
 Job time : 3123 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 05:32:17 ; Search time 4907 Seconds  
(without alignments)  
11076.455 Million cell updates/sec

Title: US-09-600-714-41  
Perfect score: 1254  
Sequence: 1 atgagctctaagtaccgcgcg.....atttggctgtggatttttaa 1254

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl :

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rat.\*

36: em.htg.wam.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1254	100.0	1254	6	AX022514	AX022514 Sequence
2	1254	100.0	1254	6	BD124102	BD124102 Novel nuc
3	1252.4	99.9	1254	9	AB018969	AB018969 Homo sapi
4	1252.4	99.9	1354	9	HUMRHDPANTI	L08429 Human Rhd b
5	1252.4	99.9	1545	9	HSRHIII	X63094 H.sapiens m
6	1252.4	99.9	2790	9	HSRHIII	X63097 H.sapiens m
7	1250.8	99.7	1254	9	AB018966	AB018966 Homo sapi
8	1250.8	99.7	1254	9	AB018967	AB018967 Homo sapi
9	1250.8	99.7	1480	9	AF510069	AF510069 Homo sapi
10	1250.8	99.7	1480	9	AF510070	AF510070 Homo sapi
11	1247.8	99.5	1251	9	AY449381	AY449381 Homo sapi
12	1247.8	99.5	1251	9	AY449382	AY449382 Homo sapi
13	1247.8	99.5	1251	9	AY449383	AY449383 Homo sapi
14	1247.8	99.5	1251	9	AY449384	AY449384 Homo sapi
15	1247.8	99.5	1251	9	AY449385	AY449385 Homo sapi
16	1247.6	99.5	1260	9	S78509	S78509 Rhd [human]
17	1246	99.4	1254	9	AB018968	AB018968 Homo sapi
18	1246	99.4	1254	9	S82449	S82449 Homo sapien
19	1242.8	99.1	1458	9	AF187846	AF187846 Homo sapi
20	1239.8	98.9	1251	9	S57971	S57971 Rh polypt
21	1239.6	98.9	1254	9	AB046420	AB046420 Homo sapi
22	1238	98.7	1328	9	AB049754	AB049754 Homo sapi
23	1231.4	98.2	1247	9	AF037626	AF037626 Homo sapi
24	1226.8	97.8	1254	9	AB037270	AB037270 Homo sapi
25	1225.2	97.7	1336	9	AF312679	AF312679 Homo sapi
26	1222	97.4	1330	9	HS297026	Z97026 Homo sapien
27	1220.4	97.3	1254	9	AF177938	AF177938 Homo sapi
28	1199.6	95.7	1328	9	AB030388	AB030388 Homo sapi
29	1199.6	95.7	1328	9	AB049753	AB049753 Homo sapi
30	1191.6	95.0	1350	9	GORRHB	L37053 Gorilla gor
31	1190	94.9	1430	9	AF510065	AF510065 Homo sapi
32	1188.4	94.8	1328	9	AB018644	AB018644 Homo sapi
33	1186.8	94.6	1350	9	CHPRHC	L37050 Pan troglod
34	1186.8	94.6	1430	9	AF510067	AF510067 Homo sapi
35	1186.8	94.6	1430	9	AF510068	AF510068 Homo sapi
36	1185.2	94.5	1384	6	I89819	I89819 Sequence 1
37	1185.2	94.5	1384	6	HUMRHPPA	M34015 Human Rh po
38	1185.2	94.5	1471	9	HSRH30A	X54534 Human mRNA
39	1183.6	94.4	1328	9	AB018645	AB018645 Homo sapi
40	1183.6	94.4	1350	9	CHPRHB	L37049 Pan troglod
41	1183.6	94.4	1430	9	AF510066	AF510066 Homo sapi
42	1182	94.3	1466	6	I89820	I89820 Sequence 2
43	1180.8	94.2	1427	9	AF056965	AF056965 Homo sapi
44	1180.6	94.1	1251	9	S57967	S57967 Rh polypt
45	1180.4	94.1	1350	9	CHPRHA	L37048 Pan troglod

# ALIGNMENTS

RESULT 1  
AX022514  
LOCUS AX022514 1254 bp DNA linear PAT 24-NOV-2000  
DEFINITION Sequence 41 from Patent WO9337763.  
ACCESSION AX022514  
VERSION AX022514.1 GI:10046112  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE  
AUTHORS Flegel, W.A. and Wagner, F.F.  
TITLE Novel nucleic acid molecules correlated with the rhesus weak d phenotype  
JOURNAL Patent: WO 9337763-A 41 29-JUL-1999;

FEATURES  
source  
CDS  
ORIGIN

100.0%; Score 1254; DB 6; Length 1254;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGCTCTAAGTACCGCGGTGTGTCGGCGCTGCTGCGCCCTCTGCGCCCTTAACTG 60  
Db  
1 ATGAGCTCTAAGTACCGCGGTGTGTCGGCGCTGCTGCGCCCTCTGCGCCCTTAACTG 60  
Qy

61 GAAGCAGCTCTCATCTCTCTCTCTATTTTACCCACTATGACGCTTCTCTAGAGAT 120  
Db  
61 GAAGCAGCTCTCATCTCTCTCTCTATTTTACCCACTATGACGCTTCTCTAGAGAT 120  
Qy

121 CAAAGGGGCTCGTGGCATCCTATCAAGTTGGCCCAAGATCTGACCGTGGCGGCATT 180  
Db  
121 CAAAGGGGCTCGTGGCATCCTATCAAGTTGGCCCAAGATCTGACCGTGGCGGCATT 180  
Qy

181 GGCTTGGGCTTCTCAGCTGAGTTTCCGAGACACAGCTGGAGCAGTGGCCCTCAAC 240  
Db  
181 GGCTTGGGCTTCTCAGCTGAGTTTCCGAGACACAGCTGGAGCAGTGGCCCTCAAC 240  
Qy

241 CTCTTCATGCTGGCGTGTGTCAGTGGCAATCTGCTGGAGCGCTTCTGAGCGCAG 300  
Db  
241 CTCTTCATGCTGGCGTGTGTCAGTGGCAATCTGCTGGAGCGCTTCTGAGCGCAG 300  
Qy

301 TTCCCTTCTGGAGGTTGTCATCACACTGTTTACGATTTGCGTGGCCCACTAGTGTCT 360  
Db  
301 TTCCCTTCTGGAGGTTGTCATCACACTGTTTACGATTTGCGTGGCCCACTAGTGTCT 360  
Qy

361 TTGTCGCTGATCTCAGTGGATGCTCTTGGGAGGTCATCTGAGGAGTGGCGAGTGGTG 420  
Db  
361 TTGTCGCTGATCTCAGTGGATGCTCTTGGGAGGTCATCTGAGGAGTGGCGAGTGGTG 420  
Qy

421 GTGATGCTGCTGGAGGTCATCACACTGTTTACGATTTGCGTGGCCCACTAGTGTATC 480  
Db  
421 GTGATGCTGCTGGAGGTCATCACACTGTTTACGATTTGCGTGGCCCACTAGTGTATC 480  
Qy

481 TTCAACACAGATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
Db  
481 TTCAACACAGATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
Qy

541 CTGTCGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600  
Db  
541 CTGTCGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600  
Qy

601 ACAGCAACGATACCCAGTTTCTGTCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 660  
Db  
601 ACAGCAACGATACCCAGTTTCTGTCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 660  
Qy

661 CCAAGTTTCACTCTGCTCTGAGAGTCCATCGAAGAGAGATGCGCTGTTCAAC 720  
Db  
661 CCAAGTTTCACTCTGCTCTGAGAGTCCATCGAAGAGAGATGCGCTGTTCAAC 720  
Qy

661 CCAAGTTTCACTCTGCTCTGAGAGTCCATCGAAGAGAGATGCGCTGTTCAAC 720  
Qy  
721 ACCTACTATGCTGTAGCAGTACGCTGCTGACAGCCTCTCAGGCTCATCTTGGCTCAC 780  
Db  
721 ACCTACTATGCTGTAGCAGTACGCTGCTGACAGCCTCTCAGGCTCATCTTGGCTCAC 780  
Qy  
781 CCCCAGGGAAGATCAGCAAGACTTATGTGCAAGTGGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db  
781 CCCCAGGGAAGATCAGCAAGACTTATGTGCAAGTGGCTGCTGCTGCTGCTGCTGCTGCT 840  
Qy  
841 GTGGGTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
Db  
841 GTGGGTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
Qy  
901 GCTGGGTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
Db  
901 GCTGGGTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
Qy  
961 GGGATTCCCCACAGCTCCATCATGGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
Db  
961 GGGATTCCCCACAGCTCCATCATGGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
Qy  
1021 ATCATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
Db  
1021 ATCATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
Qy  
1081 TTCAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
Db  
1081 TTCAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
Qy  
1141 GGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
Db  
1141 GGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
Qy  
1201 TATTTGATGACCAAGTTTCTGGAAGTTCCTCAATTTGCTGCTGCTGCTGCTGCTGCT 1254  
Db  
1201 TATTTGATGACCAAGTTTCTGGAAGTTCCTCAATTTGCTGCTGCTGCTGCTGCTGCT 1254  
Qy

RESULT 2  
BD124102  
LOCUS  
DEFINITION  
Novel nucleic acid molecule correlating to Rhesus weak D phenotype.  
ACCESSION  
BD124102  
VERSION  
BD124102.1 GI:23219047  
KEYWORDS  
JP 2002500884-A/41.  
SOURCE  
unidentified  
ORGANISM  
unclassified.  
REFERENCE  
1 (bases 1 to 1254)  
Fregel, V.A. and Wagner, F.F.  
AUTHORS  
Novel nucleic acid molecule correlating to Rhesus weak D phenotype  
TITLE  
Patent: JP 2002500884-A 41 15-JAN-2002;  
JOURNAL  
DRK BLUTSPENDEDIENST BADEN WUERTEMBERG GGBMB  
COMMENT  
OS Unidentified  
PN JP 2002500884-A/41  
PD 15-JAN-2002  
PF 18-DEC-1998 JP 2000528671  
PR 23-JAN-1998 EP 98101203.2  
PI VILLY A FREGEL, FRANZ F WAGNER  
PC  
C12N15/09, C07K14/47, C07K16/19, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10,  
PC C12P1/02, C12P21/08, C12Q1/02, C12Q1/68, G01N33/566, C12N15/00, PC  
C12N5/00  
CC Strandedness: Single;  
CC Topology: Linear;  
CC Novel nucleic acid molecule correlating to Rhesus weak D CC  
phenotype  
FH Key Location/Qualifiers  
FT CDS Location/Qualifiers  
1. .1254



```
/organism="unidentified"  
/mol_type="genomic DNA"  
/db xref="taxon:32644"
```

## ORIGIN

Query Match									
Best local Similarity 100.0%; Score 1254; DB 6; Length 1254;									
Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ATGAGCTCTAAGTACCCGCGTCTGTCGGCGCTGCTGCCCTCTGGGCCCTTAACACTG	60						
Db	1	ATGAGCTCTAAGTACCCGCGTCTGTCGGCGCTGCTGCCCTCTGGGCCCTTAACACTG	60						
Qy	61	GAAGCAGCTCTCAATCTCCTCTCTATTTTTTTATCCCACTATGACGCTTCCTTTAGAGAT	120						
Db	61	GAAGCAGCTCTCAATCTCCTCTCTATTTTTTTATCCCACTATGACGCTTCCTTTAGAGAT	120						
Qy	121	CAAAAGGGGCTCGTGGCATCTCATCAAGTTGGCCAAAGATCTGACCGTGATGGCGGCATT	180						
Db	121	CAAAAGGGGCTCGTGGCATCTCATCAAGTTGGCCAAAGATCTGACCGTGATGGCGGCATT	180						
Qy	181	GGCTTGGGCTTCCTCACCTCGAGTTTCGGAGACACAGCTGGAGCAGTGGCCCTTCAAC	240						
Db	181	GGCTTGGGCTTCCTCACCTCGAGTTTCGGAGACACAGCTGGAGCAGTGGCCCTTCAAC	240						
Qy	241	CTCTTCTATGCTGGCGCTGTGTGTGCAGTGGGCAATCCTGTGGACGGCTTCCTGAGCCAG	300						
Db	241	CTCTTCTATGCTGGCGCTGTGTGTGCAGTGGGCAATCCTGTGGACGGCTTCCTGAGCCAG	300						
Qy	301	TTCCCTTCTCGGAAGTGCTCATCACACTGTTCAAGTATTCGGCTGGGCCACCATGAGTGCT	360						
Db	301	TTCCCTTCTCGGAAGTGCTCATCACACTGTTCAAGTATTCGGCTGGGCCACCATGAGTGCT	360						
Qy	361	TTGTTCGGTCTGATCTCAGTGGATGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTTCGGTG	420						
Db	361	TTGTTCGGTCTGATCTCAGTGGATGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTTCGGTG	420						
Qy	421	GTGATGCTGCTGGTGGAGGTGACAGCTTTTAGGCAACTGAGGATGGTTCATCAGTAATATC	480						
Db	421	GTGATGCTGCTGGTGGAGGTGACAGCTTTTAGGCAACTGAGGATGGTTCATCAGTAATATC	480						
Qy	481	TTCAACACAGACTACCAATGAACTATCATGATGCACATCTACCTGTTCGCAGCCTATTTTGGG	540						
Db	481	TTCAACACAGACTACCAATGAACTATCATGATGCACATCTACCTGTTCGCAGCCTATTTTGGG	540						
Qy	541	CTGTCTGTGGCTGGTGCCTGCCAAAGCCCTTACCCGAGGGAACGGAGGATAAAGATCAG	600						
Db	541	CTGTCTGTGGCTGGTGCCTGCCAAAGCCCTTACCCGAGGGAACGGAGGATAAAGATCAG	600						
Qy	601	ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGGGCCCTCTTCTTGTGGATGTTCTGG	660						
Db	601	ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGGGCCCTCTTCTTGTGGATGTTCTGG	660						
Qy	661	CCAAGTTTCAACTCTGCTCTGCTGAGAAGTCCAACTCGAAAGGAAGATCCGCTGTTCAAC	720						
Db	661	CCAAGTTTCAACTCTGCTCTGCTGAGAAGTCCAACTCGAAAGGAAGATCCGCTGTTCAAC	720						
Qy	721	ACCTACTATGCTGTAGCAGTACAGGTGGTGACAGCCATCTCAGGGTCACTCCTTTGGCTCAC	780						
Db	721	ACCTACTATGCTGTAGCAGTACAGGTGGTGACAGCCATCTCAGGGTCACTCCTTTGGCTCAC	780						
Qy	781	CCCCAAGGGAAGATCAGCAAGCTTATGTGCACAGTGCAGTGTGGCAGAGAGCGTGGCT	840						
Db	781	CCCCAAGGGAAGATCAGCAAGCTTATGTGCACAGTGTGGCAGAGAGCGTGGCT	840						
Qy	841	GTGGGTACCTCGTGTCACTGATCCCTTCTCCGTGGCTTGCCTATGGTGTCTGGGCTTGTG	900						
Db	841	GTGGGTACCTCGTGTCACTGATCCCTTCTCCGTGGCTTGCCTATGGTGTCTGGGCTTGTG	900						
Qy	901	GCTTGGGCTGATCTCCGTCGGGGAGCCAGTACCTGCGGGGTGTGTATACCGAGTGTG	960						
Db	901	GCTTGGGCTGATCTCCGTCGGGGAGCCAGTACCTGCGGGGTGTGTATACCGAGTGTG	960						

Qy	961	GGGATTC	CCCA	CAGCTCC	ATCAT	GGGCTCA	CAACTT	CAGCTT	GTGGGT	CTGCTT	GGAGAG	1020	
Db	961	GGGATTC	CCCA	CAGCTCC	ATCAT	GGGCTCA	CAACTT	CAGCTT	GTGGGT	CTGCTT	GGAGAG	1020	
Qy	1021	ATCATCT	ACAT	TGTGCT	GTGGT	CTTGAT	ATACCGT	TCGGAGC	CGGCAAT	GCATCAT	TGGC	1080	
Db	1021	ATCATCT	ACAT	TGTGCT	GTGGT	CTTGAT	ATACCGT	TCGGAGC	CGGCAAT	GCATCAT	TGGC	1080	
Qy	1081	TTCAGGT	CTCT	CTCAG	CAAT	TGGGA	CACTCAG	CTTG	GCCAT	CTGTAT	GTCTC	ACAGTCT	1140
Db	1081	TTCAGGT	CTCT	CTCAG	CAAT	TGGGA	CACTCAG	CTTG	GCCAT	CTGTAT	GTCTC	ACAGTCT	1140
Qy	1141	GGTCT	CTCT	GACAG	GTGCT	CTCTAA	TCTTAA	ATAT	TGGAAC	CACTCT	CATGAGG	CTAAA	1200
Db	1141	GGTCT	CTCT	GACAG	GTGCT	CTCTAA	TCTTAA	ATAT	TGGAAC	CACTCT	CATGAGG	CTAAA	1200
Qy	1201	TATTTTG	ATGAC	CAAG	TTTTCT	GGAA	GTTCCT	CATT	TGGCT	TTGGAT	TTTTAA	1254	
Db	1201	TATTTTG	ATGAC	CAAG	TTTTCT	GGAA	GTTCCT	CATT	TGGCT	TTGGAT	TTTTAA	1254	

### RESULT 3

RESULTS	AB018969	1254 bp	linear	PRI 03-JUL-1999
LOCUS	Homo sapiens RHDY0 mRNA for Rh blood group D antigen (RHD), complete cds.			
DEFINITION	AB018969			
ACCESSION	AB018969, 1 GI:5360898			
VERSION	RHDY0; Rh blood group D antigen (RHD).			
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (sites)			
AUTHORS	Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y., Tsuneyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.			
TITLE	Polymorphisms of RhdVa in Japanese			
REFERENCE	2 (bases 1 to 1254)			
AUTHORS	Uchikawa, M., Hyodo, H. and Ishikawa, Y.			
TITLE	Direct Submision			
JOURNAL	Submitted (20-OCT-1998) Hironobu Hyodo, Japanese Red Cross Central Blood Center, Research; 4-1-31, Hiroo, Shibuya, Tokyo 150-0012, Japan (E-mail:hyodo@hla.cbc.jrc.or.jp, Tel: +81-3-5465-6009, Fax: +81-3-3406-7832)			

## FEATURES source

```

1. 11254
/organism="Homo sapiens"
/mol_type="mRNA"
/isolate="YO"
/db_xref="taxon:9606"
1. 11254
/gene="RhDYO"
1. 1254
/gene="RhDYO"
/codon_start=1
/note="RhD variant RhDYO"
/product="Rh blood group D antigen (RHD)"
/protein_id="BAA82159.1"
/db_xref="GI:5360899"
/translation="MSSKPKRSVRCLPLMALTLAAALLLFYFETHYDASLEOKGL
VASQVGGDLTWMAIGIGELTSSPRHSWSVAENFMALAGVQWAILIDGTFVQCFP
SGKVYITLFTSLATMSALSVLISDVADPKGVKNTLAQLVVMVLVEVTAIGNLRWVLSIN
FNIDVHNNMHVIFYFAAYFGVLSVADCLPKPIEGESTKDOTATPSLSAMLGALFLWM
FWPNSCHLLASPIKRNKNAVNTYYAVASVVVTAISGSLAHPOCKISKTYVHGAVLA
GGVAVGTSCHLITPSPLWAMVLVAGLISVGCAKVLPGCCNRVLGIPHSSIMGYNFSL
LGLLGEIIVYVLLVDTVGAGNGMIGFQVLLSIGELSLAIVIAITSGELTGLLLNLIK
WKAPAEAKYFDDDPKVPFHLAVGP"

```

## ORIGIN

```
Query Match          99.9%; Score 1252.4; DB 9; Length 1254;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGAGCTCTAAGTACCGCGGTCGTCCCGCGTGTGCTGCCCTCTGCGGCCCTTAAACACTG 60
Db 1 ATGAGCTCTAAGTACCGCGGTCGTCCCGCGTGTGCTGCCCTCTGCGGCCCTTAAACACTG 60
QY 61 GAGCAGCTCTCATCTCCCTCTCTATTTTACCCACTATGACGCTTCTTAGAGGAT 120
Db 61 GAGCAGCTCTCATCTCCCTCTCTATTTTACCCACTATGACGCTTCTTAGAGGAT 120
QY 121 CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAAGATCGACCGTGATGGCGGCATT 180
Db 121 CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAAGATCGACCGTGATGGCGGCATT 180
QY 181 GGCTTGGCTTCTCTACCTCGAGTTCCGGAGACACAGCTGGAGCATGTGGCTTCAAC 240
Db 181 GGCTTGGCTTCTCTACCTCGAGTTCCGGAGACACAGCTGGAGCATGTGGCTTCAAC 240
QY 241 CTCTTCATGCTGGCGTGTGGTGGCATCTGAGTGGGCAATCTGCTGGACGCTTCTTGAGCCAG 300
Db 241 CTCTTCATGCTGGCGTGTGGTGGCATCTGAGTGGGCAATCTGCTGGACGCTTCTTGAGCCAG 300
QY 301 TTCCTTCTGGAGAGTGGTCATCACACTGTTTCAGTATTCGGCTGGCCACCATGAGTGT 360
Db 301 TTCCTTCTGGAGAGTGGTCATCACACTGTTTCAGTATTCGGCTGGCCACCATGAGTGT 360
QY 361 TTGTGCGTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGT 420
Db 361 TTGTGCGTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGT 420
QY 421 GTGATGTGTGGTGGAGTGCACAGCTTTAGGCAACCTGAGATGCTCATCAGTAATATC 480
Db 421 GTGATGTGTGGTGGAGTGCACAGCTTTAGGCAACCTGAGATGCTCATCAGTAATATC 480
QY 481 TTCAACACAGACTACCATGAACATGATGACATCTACGTGTTTCGAGGCTATTTGGG 540
Db 481 TTCAACACAGACTACCATGAACATGATGACATCTACGTGTTTCGAGGCTATTTGGG 540
QY 541 CTGTCTGTGGCTGGTGCCTCCAAAGCTTACCCGAGGAAACGGAGGATAAGATCAG 600
Db 541 CTGTCTGTGGCTGGTGCCTCCAAAGCTTACCCGAGGAAACGGAGGATAAGATCAG 600
QY 601 ACAGCAAGTACCCAGTTTCTGCTGCCATGTGGCGCCCTCTTCTTGTGATGTTCTGG 660
Db 601 ACAGCAAGTACCCAGTTTCTGCTGCCATGTGGCGCCCTCTTCTTGTGATGTTCTGG 660
QY 661 CCAAGTTTCAACTCTGCTCTGAGAGTCCAATCGAAGAGAAATGCCGTGTTCAAC 720
Db 661 CCAAGTTTCAACTCTGCTCTGAGAGTCCAATCGAAGAGAAATGCCGTGTTCAAC 720
QY 721 ACCTACTATGCTGAGCAGTCCAGTGTGACAGCCATCTCAGGCTCATCTTGGCTCAC 780
Db 721 ACCTACTATGCTGAGCAGTCCAGTGTGACAGCCATCTCAGGCTCATCTTGGCTCAC 780
QY 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCACAGTGGTGTGGCAGGAGCGGTGCT 840
Db 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCACAGTGGTGTGGCAGGAGCGGTGCT 840
QY 841 GTGGGTACCTCGTGTCACTGATTCCTTCTCGTGGCTTGCATGGTCTGCGTCTTGTG 900
Db 841 GTGGGTACCTCGTGTCACTGATTCCTTCTCGTGGCTTGCATGGTCTGCGTCTTGTG 900
QY 901 GCTGGGTGATCTCGTGGGGAGCCAAAGTACCTGCGGGGTGTGTAAACCGAGTGTG 960
Db 901 GCTGGGTGATCTCGTGGGGAGCCAAAGTACCTGCGGGGTGTGTAAACCGAGTGTG 960
QY 961 GGGATTTCCCAAGCTCCATCATGGGCTTACACTTTCAGTCTGCTGGGTCTGTTGAGAG 1020
Db 961 GGGATTTCCCAAGCTCCATCATGGGCTTACACTTTCAGTCTGCTGGGTCTGTTGAGAG 1020
QY 1021 ATCATCTACATGCTGCTGCTGATGATACCGTTCGAGCGGCAATGGCATGATTCGC 1080
Db 1021 ATCATCTACATGCTGCTGCTGATGATACCGTTCGAGCGGCAATGGCATGATTCGC 1080
```

```
QY 1081 TTCCAGGTCCTCTCAGCATTTGGGAACCTCAGCTTGGCCATCGCTAGTACTCTCAGCTCT 1140
Db 1081 TTCCAGGTCCTCTCAGCATTTGGGAACCTCAGCTTGGCCATCGCTAGTACTCTCAGCTCT 1140
QY 1141 GGTCTCTCTCAGCAGTTTGTCTCTAAATCTTAAATATGGAACACCTCATGAGGCTAAA 1200
Db 1141 GGTCTCTCTCAGCAGTTTGTCTCTAAATCTTAAATATGGAACACCTCATGAGGCTAAA 1200
QY 1201 TATTTTGTGACCAAGTTTCTGGAAGTTTCCTCATTTGGCTGTTGGATTTAA 1254
Db 1201 TATTTTGTGACCAAGTTTCTGGAAGTTTCCTCATTTGGCTGTTGGATTTAA 1254
```

RESULT 4

HUMRHDAINTI

LOCUS HUMRHDAINTI 1354 bp mRNA linear PRI 24-AUG-1993

DEFINITION Human RHD blood group antigen mRNA, complete cds.

ACCESSION L08429

VERSION 1

KEYWORDS Rhd blood group; Rhd blood group; antigen; blood group antigen.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1354)

AUTHORS Arce, M.A., Thompson, E.S., Wagner, S., Coyne, K.E., Ferdman, B.A. and Lublin, D.M.

TITLE Molecular cloning of Rhd cDNA derived from a gene present in RHD-positive, but not RHD-negative individuals

JOURNAL Blood 82 (2), 651-655 (1993)

MEDLINE 93320449

PUBMED 8329718

COMMENT Original source text: Homo sapiens (library: HL1058b (from Clontech)) Bone marrow cDNA to mRNA.

FEATURES

source

1..1354

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

1..1254

/standard\_name="Rhd"

/citations=[1]

/codon\_start=1

/evidence=experimental

/product="Rhd blood group antigen"

/protein\_id="AAA02679.1"

/db\_xref="GI:337391"

/translation="MSKYPKRRRCPLMALTLLEALILLFFHYTHDASLEQKGL VASYVQGDITVMAAIGLGLTSPRRHSVSNFNLMLALGVQWAILLDGFIQFP SGKVIITLIRLATMSALSVLISVDAVLGKVLNLAQVVMVLVEVTLGNLRLVSNL FNTDYMNMWHIYVFAAYFGLSVAMCLPKPEGTEDKDTATIPSLSAMLGALFLWM FWPFSNALLRSPIERKNVNTYVAVSVVTAISGSLAHPOGKISKTVHSAVLA GGVAVTSGSLIIPSPMLNLVAGLISVGGAKYLPCCNRVLGIPHSSIMGVNFSL LGDLGEIYIIVLLDITVGAGNGMIGFQVLLSISGELSIAIVALTSGLLTGLLLNLKI WKAPHEAKYFDDQVWFKPHLVGFG"

CDS

Query Match 99.9%; Score 1252.4; DB 9; Length 1354;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGCTCTAAGTACCGCGGTCGTCCCGCGTGTGCTGCCCTCTGCGGCCCTTAAACACTG 60

Db 1 ATGAGCTCTAAGTACCGCGGTCGTCCCGCGTGTGCTGCCCTCTGCGGCCCTTAAACACTG 60

QY 61 GAGCAGCTCTCATCTCCCTCTCTATTTTACCCACTATGACGCTTCTTAGAGGAT 120

Db 61 GAGCAGCTCTCATCTCCCTCTCTATTTTACCCACTATGACGCTTCTTAGAGGAT 120

QY 121 CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAAGATCGACCGTGATGGCGGCATT 180

Db 121 CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAAGATCGACCGTGATGGCGGCATT 180







Qy	181	GGCTTGGGCTTCCTCACCTCGAGTTTCCGAGACACAGCTGGAGCAGTGTGGCCCTCAAC	240
Db	181	GGCTTGGGCTTCCTCACCTCGAGTTTCCGAGACACAGCTGGAGCAGTGTGGCCCTCAAC	240
Qy	241	CTCTTCATGCTGGCGCTTGGTGGCACTGCTGGAGGCAATCTCTGTCAGACGGCTTCCTGAGCCAG	300
Db	241	CTCTTCATGCTGGCGCTTGGTGGCACTGCTGCTGAGCGGCTTCCTGAGCCAG	300
Qy	301	TTCCCTTTCTGGAAAGTGGTCAACACACTGTTTCTGATATTCGGCTGGCCACCAGTAGTGTCT	360
Db	301	TTCCCTTTCTGGAAAGTGGTCAACACACTGTTTCTGATATTCGGCTGGCCACCAGTAGTGTCT	360
Qy	361	TTGTGCGTGTGATCTCAGTGGATGCTGTTCTGGGGAAGTCAACTTGGCGCAGTTTGGTG	420
Db	361	TTGTGCGTGTGATCTCAGTGGATGCTGTTCTGGGGAAGTCAACTTGGCGCAGTTTGGTG	420
Qy	421	GTGATGTGCTGTGGAGGTGACAGCTTTTAGGCAACCTGAGGATGTTTCATCAGTAATATC	480
Db	421	GTGATGTGCTGTGGAGGTGACAGCTTTTAGGCAACCTGAGGATGTTTCATCAGTAATATC	480
Qy	481	TTCAACACAGACTACCCACATGAACATGATGCACATCTACGTTTCCGAGCCTATTTTGGG	540
Db	481	TTCAACACAGACTACCCACATGAACATGATGCACATCTACGTTTCCGAGCCTATTTTGGG	540
Qy	541	CTGTCTGTGGCTGTGCTGCCAAAGCTCTACCCGAGGGAACGGAGGATAAAGATCAG	600
Db	541	CTGTCTGTGGCTGTGCTGCCAAAGCTCTACCCGAGGGAACGGAGGATAAAGATCAG	600
Qy	601	ACAGCAAGATACCCAGTTTGTCTGCGAATGCTGGGCGCCCTCTTCTTTGTGGATGTTCTGG	660
Db	601	ACAGCAAGATACCCAGTTTGTCTGCCATGCTGGGCGCCCTCTTCTTTGTGGATGTTCTGG	660
Qy	661	CCAAAGTTTCAACTGTCTCTCAGAGAGTCCAAATCGAAAGGAAGATCCGCTGTTCAAC	720
Db	661	CCAAAGTTTCAACTGTCTCTCAGAGAGTCCAAATCGAAAGGAAGATCCGCTGTTCAAC	720
Qy	721	ACCTACTATGCTGTAGCAGTCCAGCGTGGTGACGCCATCTCAGGGTCATCCTTGGCTCAC	780
Db	721	ACCTACTATGCTGTAGCAGTCCAGCGTGGTGACGCCATCTCAGGGTCATCCTTGGCTCAC	780
Qy	781	CCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGGGTGTTGGCAGAGCGGTGGCT	840
Db	781	CCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGGGTGTTGGCAGAGCGGTGGCT	840
Qy	841	GTGGGTACCTCGTGTCACTGATCCCTTCTCCGTGGGCTTGCCATGCTGCTGGGCTTGTG	900
Db	841	GTGGGTACCTCGTGTCACTGATCCCTTCTCCGTGGGCTTGCCATGCTGCTGGGCTTGTG	900
Qy	901	GCTGGGCTGATCTCCGTGGGGAGCCAAAGTACCTGCGGGGGTGTGTAACCGAGTCTG	960
Db	901	GCTGGGCTGATCTCCGTGGGGAGCCAAAGTACCTGCGGGGGTGTGTAACCGAGTCTG	960
Qy	961	GGGATTTCCCAACAGCTCCATCATGGGCTTACAACCTTCAGTCTGCTGGGCTGCTGGAGAG	1020
Db	961	GGGATTTCCCAACAGCTCCATCATGGGCTTACAACCTTCAGTCTGCTGGGCTGCTGGAGAG	1020
Qy	1021	ATCATCTACATTTGCTGCTGGTGTTCGATACCGTCGAGCGCCGAATGGCATGATTTGCC	1080
Db	1021	ATCATCTACATTTGCTGCTGGTGTTCGATACCGTCGAGCGCCGAATGGCATGATTTGCC	1080
Qy	1081	TTCCAGGTCCTCCTCAGCATTTGGGAACTCAGTCTGGCCATCTGATAGTCTCTCACGCTCT	1140
Db	1081	TTCCAGGTCCTCCTCAGCATTTGGGAACTCAGTCTGGCCATCTGATAGTCTCTCACGCTCT	1140
Qy	1141	GGTCTCCTGACAGTTTGTCTCTAAATCTTAAATATGGAAGCACTTCATGAGGCTTAA	1200
Db	1141	GGTCTCCTGACAGTTTGTCTCTAAATCTTAAATATGGAAGCACTTCATGAGGCTTAA	1200
Qy	1201	TATTTTGTATGACCAAGTTTCTGGAGTTTCTCATTTGGCTGTGGATTTTAA	1254
Db	1201	TATTTTGTATGACCAAGTTTCTGGAGTTTCTCATTTGGCTGTGGATTTTAA	1254

RESULT 8	AB018967	1254 bp	mRNA	linear	PRI 03-JUL-1999
LOCUS	Homo sapiens RhdVa(TO)	1254 bp	mRNA	linear	PRI 03-JUL-1999
DEFINITION	Homo sapiens RhdVa(TO) mRNA for Rh blood group D antigen (RHD), complete cds.				
ACCESSION	AB018967				
VERSION	AB018967.1	GI:5360244			
KEYWORDS	RHDVa(TO); Rh blood group D antigen (RHD).				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (sites)				
AUTHORS	Hyodo.H., Ishikawa.Y., Kashiwase.K., Ogawa.A., Watanabe.Y., Tsuneyama.H., Toyoda.C., Uchikawa.M., Akaza.T. and Fujii.T.				
JOURNAL	Polymorphisms of RhdVa in Japanese				
REFERENCE	Unpublished				
AUTHORS	2 (bases 1 to 1254)				
TITLE	Uchikawa.M., Hyodo.H. and Ishikawa.Y.				
AUTHORS	Direct Submission				
TITLE	Submitted (20-OCT-1998) Hironchu Hyodo, Japanese Red Cross Central Blood Center, Research; 4-1-31, Hiroco, Shibuya, Tokyo 150-0012,				
JOURNAL	Japan (E-mail:hyodo@hia.cbc.or.jp, Tel:+81-3-5485-6009, Fax:+81-3-3406-7892)				
FEATURES	Location/Qualifiers				
source	1..1254				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/isolate="TO"				
	/db_xref="taxon:9606"				
gene	1..1254				
	/gene="RhdVa(TO)"				
CDS	1..1254				
	/gene="RhdVa(TO)"				
	/note="Rh variant RhdVa(TO)"				
	/codon_start=1				
	/product="Rh blood group D antigen (RHD)"				
	/protein_id="BAA81900.1"				
	/db_xref="GI:5360245"				
	/translation="MSSKIPSVRRCLPLNALTLEAMILLIFYFTHVDASLDQKGI VASVQVQDITVMAAIGLGFSTFRHSSVAFNFMALGVQWAILLDGFLSQPFI SKVYITLFLRIATMGALSIVSDAVLGKVNLAQLVWVLVEVTLGNLRMVISNNI PNTDYNNMMHVIYVFAAVFGLSVAMCLPKLPPECTEDKDOTATIPSLAMGLAFMLN PWFNSALLRSPQRKNAMNTYVAVSVVTAISGSSLAHPQKISKTVYHSAVLA GGVATVSHIPLSPFLMAMVLGLVAGLISVGGAAYLPGCCNRLVGIPIHSSIMGVNFSI LGLGEIYIVLLVLDITVGAGNGMIGFQVLLISIGELSLAIVALTSGLLTGLLLNLK WKAPHEAKYFDQVFWKFPPLAVG"				
ORIGIN					
Query Match	99.7%;	Score 1250.8;	DB 9;	Length 1254;	
Best Local Similarity	99.8%;	Pred. No. 0;			
Matches 1252;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
QY	1	ATGAGCTCTAAGTACCCGCGTCTGTCGCGCGCTGCCTGCCCTCTGGGCCCTAACACTG	60		
Db	1	ATGAGCTCTAAGTACCCGCGTCTGTCGCGCGCTGCCTGCCCTCTGGGCCCTAACACTG	60		
QY	61	GAAGCAGCTCTCAATCTCCTCTCTTATTTTTTACCCACTATGAGCTTCTCTTAGAGAT	120		
Db	61	GAAGCAGCTCTCAATCTCCTCTCTTATTTTTTACCCACTATGAGCTTCTCTTAGAGAT	120		
QY	121	CAAAAGGGCTCGTGGCATCTATCAAGTTCGCCAAGATCTGACCGTGATCGCGCCATT	180		
Db	121	CAAAAGGGGCTCGTGGCATCTATCAAGTTCGCCAAGATCTGACCGTGATCGCGCCATT	180		
QY	181	GGCTTGGGCTTCTCACTCGAGTTTCGGGACACACAGCTGGAGCAGTGTGCCTTCAAC	240		
Db	181	GGCTTGGGCTTCTCACTCGAGTTTCGGGACACACAGCTGGAGCAGTGTGCCTTCAAC	240		
QY	241	CTCTTCATGCTGGCGCTTGGTGTGAGTGGGCAATCTCTGTGAGCGGCTTCTGAGCCAG	300		
Db	241	CTCTTCATGCTGGCGCTTGGTGTGAGTGGGCAATCTCTGTGAGCGGCTTCTGAGCCAG	300		

```
QY 301 TTCCCTCTCGGAGGTGGTTCATCACACTGTTTCAGTATTCGGCTGGCCACCATGAGTCT 360
Db 301 TTCCCTCTCGGAGGTGGTTCATCACACTGTTTCAGTATTCGGCTGGCCACCATGAGTCT 360
QY 361 TTGTCGGTGTGATCTCAGTGGATGCTGCTTGGGGAAGTCAACTTGGCGCATGTTGGTG 420
Db 361 TTGTCGGTGTGATCTCAGTGGATGCTGCTTGGGGAAGTCAACTTGGCGCATGTTGGTG 420
QY 421 GTGATGCTGCTGGAGGTGACAGCTTTAGGCAACCTGAGGATGTCATCAGTAATATC 480
Db 421 GTGATGCTGCTGGAGGTGACAGCTTTAGGCAACCTGAGGATGTCATCAGTAATATC 480
QY 481 TTCAACACAGACTACCACTGAATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 481 TTCAACACAGACTACCACTGAATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 CTGCTGTGCTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 CTGCTGTGCTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 ACAGCAACGATACCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 ACAGCAACGATACCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 CCAAGTTTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 CCAAGTTTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 ACTTACTGCTGTAGAGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780
Db 721 ACTTACTGCTGTAGAGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780
QY 781 CCCAAGGGAAGATCAGCAAGCTTATGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 840
Db 781 CCCAAGGGAAGATCAGCAAGCTTATGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 840
QY 841 GTGGGTACCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 841 GTGGGTACCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 901 GTGGGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 GTGGGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 GGGATTCCTCAGAGTCCATCATGAGTCCATCATGAGTCCATCATGAGTCCATCATGAG 1020
Db 961 GGGATTCCTCAGAGTCCATCATGAGTCCATCATGAGTCCATCATGAGTCCATCATGAG 1020
QY 1021 ATCATCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 ATCATCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1081 TTCCAGTCTCTCAGAGTGGGAACTCAGTGGGAACTCAGTGGGAACTCAGTGGGAACTC 1140
Db 1081 TTCCAGTCTCTCAGAGTGGGAACTCAGTGGGAACTCAGTGGGAACTCAGTGGGAACTC 1140
QY 1141 GGTCTCTCAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1200
Db 1141 GGTCTCTCAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1200
QY 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTGGAAGTTTCTGGAAGTTTCTGGAAGTT 1254
Db 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTGGAAGTTTCTGGAAGTTTCTGGAAGTT 1254
```

## RESULT 9

AF510069

LOCUS

DEFINITION Homo sapiens Rhesus blood group D antigen D(667) variant mRNA, complete cds.

ACCESSION AF510069

VERSION AF510069.1

GI:25573107

KEYWORDS

## SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1480)  
Noizat-Pirenne, F., Lee, K., Le Pennec, P. Y., Simon, P., Kazup, P.,  
Bachir, D., Rouzaud, A. M., Roussel, M., Juscak, G., Menanteau, C.,  
Rouger, P., Kobb, R., Carttron, J. P. and Ansart-Pirenne, H.  
Rare RHCE phenotypes in black individuals of Afro-Caribbean origin:  
identification and transfusion safety  
Blood 100 (12), 4223-4231 (2002)  
12393640  
2 (bases 1 to 1480)  
Noizat-Pirenne, F., Carttron, J. P. and Ansart-Pirenne, H.  
Direct Submission  
Submitted (08-MAY-2002) CNRGS, INTS, 20 rue Bouvier, Paris 75011,  
France

## Location/Qualifiers

1. .1480

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/note="isolated from Black individual of Afro-Caribbean

origin"

44. .1297

/notes="RHD"

/codon\_start=1

/product="Rhesus blood group D antigen D(667) variant"

/protein\_id="AA075125.1"

/db\_xref="GI:25573108"

/translation="MSSKYPKRSVRCLPLWALTLLEAALLLFFVFFHYDASLEDQKGL

VASYQVQDLTMAALIGLFTLSFRRHSWVAENFLMVLGVQVAILDGLDGLSFPF

SKVLTLSIRLATMSALSVLISVDALGVKNLAQVVMVLVEVLRNLRVLSNI

FNTDVMNMHHIYFAAYPGLSVAKLPKPEPTEDKDTATIPSLAMGLAFJWI

FPPNSALLRSPFERKNAFNTYAVAVSVVTAI SSSLAHPQGIKSTYYRSVAVIA

GGVAVTGSCHLIPSPWLAIVGLVAGLSVGAKYLPGCCNRVLGPHSSIMGNPSL

LGLGELIYIVLLVLDTVGAGNGMIGFVLLSIGLSLAIVLALTSGLLTGLLLNLKI

WKAPREAKYFDQVPKPFPLAVGP"

Query Match 99.7%; Score 1250.8; DB 9; Length 1480;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTCTAAGTACCCGCGGTCTGTCGCGCGGTGCTGCCCTCTGGGCCCTTAACATG 60

Db 44 ATGAGCTCTAAGTACCCGCGGTCTGTCGCGCGGTGCTGCCCTCTGGGCCCTTAACATG 103

QY 61 GAAGCAGCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120

Db 104 GAAGCAGCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 163

QY 121 CAAAAGGGCTCTGGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180

Db 164 CAAAAGGGCTCTGGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 223

QY 181 GCGTTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240

Db 224 GCGTTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 283

QY 241 CTCCTTCATGCTGGCGGTGCTGTCGAGTGGGCAATCTCTGAGGCTTCTCTGAGCCAG 300

Db 284 CTCCTTCATGCTGGCGGTGCTGTCGAGTGGGCAATCTCTGAGGCTTCTCTGAGCCAG 343

QY 301 TTCCCTTCTGGGAAGTGGTTCATCACACTGTTTCAGTATTCGGCTGGCCACCATGAGTGTCT 360

Db 344 TTCCCTTCTGGGAAGTGGTTCATCACACTGTTTCAGTATTCGGCTGGCCACCATGAGTGTCT 403

QY 361 TTGTCGGTGTGATCTCAGTGGATGCTGCTTTGGGGAAGGTCAACTGCGCAGTTGGTG 420

Db 404 TTGTCGGTGTGATCTCAGTGGATGCTGCTTTGGGGAAGGTCAACTGCGCAGTTGGTG 463

QY 421 GTGATGGTGTCTGGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGTCATCAGTAATATC 480







```
Db 584 CTGTCCTGTGGCGCTGTGCTGCCAAGCCTCTACCCGAGGGAACGAGGATAAAGATCAG 643
Qy 601 ACAGCAACGATACCCAGTTTGTCTCCATGCTGGCGCCCTCTTCTTGTGGATGTTCTGG 660
Db 644 ACAGCAACGATACCCAGTTTGTCTCCATGCTGGCGCCCTCTTCTTGTGGATGTTCTGG 703
Qy 661 CCAAGTTTCAACTGCTCTGTGTGAGAGTCAANTGAAAGAAAGAAAGCCGTGTTCAAC 720
Db 704 CCAAGTTTCAACTTGTCTGTGTGAGAGTCAANTGAAAGAAAGAAAGCCGTGTTCAAC 763
Qy 721 ACCTACTATGCTGTAGCAGTCAAGCGTGTGTGACAGCATCTCAGGGTCACTCTTGGCTCAC 780
Db 764 ACCTACTATGCTGTAGCAGTCAAGCGTGTGTGACAGCATCTCAGGGTCACTCTTGGCTCAC 823
Qy 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCAAGTGTGCGGCTTTCGAGGAGCGGTGGCT 840
Db 824 CCCCAGGGAAGATCAGCAAGACTTATGTGCAAGTGTGCGGCTTTCGAGGAGCGGTGGCT 883
Qy 841 GTGGGTACCTGTGTCACTGATCCCTTCTCGTGGCTTGCCATGGTGGGTCTTGTG 900
Db 884 GTGGGTACCTGTGTCACTGATCCCTTCTCGTGGCTTGGCCATGGTGGGTCTTGTG 943
Qy 901 GCTGGGCTGATCTCCGTCCGGGGAGCCAAAGTACCTGCGGGGTGTGTAAACGAGTGTG 960
Db 944 GCTGGGCTGATCTCCGTCCGGGGAGCCAAAGTACCTGCGGGGTGTGTAAACGAGTGTG 1003
Qy 961 GGGATTCGCCACAGCTCCATCATGCGGTACAACTTCAGCTTGTGGGTCTGTTGGAGAG 1020
Db 1004 GGGATTCGCCACAGCTCCATCATGCGGTACAACTTCAGCTTGTGGGTCTGTTGGAGAG 1063
Qy 1021 ATCATCTACATTTGCTGTGTGCTGTGTGATACCTGCGGAGCGGCAATGGCATGATTGGC 1080
Db 1064 ATCATCTACATTTGCTGTGTGCTGTGTGATACCTGCGGAGCGGCAATGGCATGATTGGC 1123
Qy 1081 TTCCAGGTCCTCTCCTCAGCAATGGGGAACCTCAGTTGGCCATGCTGATAGCTCTCAGCTCT 1140
Db 1124 TTCCAGGTCCTCTCCTCAGCAATGGGGAACCTCAGTTGGCCATGCTGATAGCTCTCAGCTCT 1183
Qy 1141 GGTCTCTCAGAGTTTGTCTCTTAATCTTAAATGAAAGACCTCATCAGGCTAA 1200
Db 1184 GGTCTCTCAGAGTTTGTCTCTTAATCTTAAATGAAAGACCTCATCAGGCTAA 1243
Qy 1201 TATTTGTATGACCAAGTTTCTGGAGTTTCTCTCAATTTGGCTGTGGATTTAA 1254
Db 1244 TATTTGTATGACCAAGTTTCTGGAGTTTCTCTCAATTTGGATTTAA 1297
```

```
RESULT 11
AY449381
LOCUS Homo sapiens rhesus blood group D antigen DFE mRNA, partial cds.
DEFINITION
ACCESSION AY449381
VERSION AY449381.1 GI:38565510
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1251)
AUTHORS Noizat-Pirenne,F. and Ansart-Pirenne,H.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2003) Blood Bank, EFS, 151 Avenue du Marechal de
Lattre de Tassigny, Creteil 94010, France
FEATURES
source
1..1251
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon.9606"
1..>1251
/notes="weak D DFE"
/codon_start=1
/product="rhesus blood group D antigen DFE"
/protein_id="AAR24078.1"
```

```
/db_xref="GI:38565511"
/translation="MSSKIPRSVRCLPLWALTLEAALLLFYFTHYDASLEDEKGL
SKVYVIGDGLTVAALIGLFTSSFRHSSVAFNLFLMFLAQLVWMLVEVTAALNGLVFLP
SGYVITLFSIRLATMSALSVLISDAVLGKVNLAQLVWMLVEVTAALNGLVFLP
PNTDYHNMHIIYVFAAYFGLSVACLPKPLPEGTEDQDQATIPSLSAMLGALFLWI
PWPNSALLRSPYERKNNAVNTYVAVASVVTATSGSSLARHPQKISKTYVHSAFLA
GDVAVGTSCHLIPSPWLAMVLGLVAGLISVGAKYLPCCNRLVIGIPHSSIMGVNPSL
LGLLGEIYYIVLLVDLTGAGNGMIGFQVLLSIGELSLAIVALTSGLLTGLLLNLKI
WKAPHEAKYFDQVFWKFFHLAVGF"
```

## ORIGIN

```
Query Match
Best Local Similarity 99.5%; Score 1247.8; DB 9; Length 1251;
Matches 1249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAGCTCTAAGTACCCGCGTGTGTCGGCGCTCCCTGCGCCCTTAACACTG 60
Db 1 ATGAGCTCTAAGTACCCGCGTGTGTCGGCGCTCCCTGCGCCCTTAACACTG 60
Qy 61 GAAGCAGCTCTCATCTCTCTCTATTTTATCCCACTATGACGCTTCTTAGAGAT 120
Db 61 GAAGCAGCTCTCATCTCTCTCTATTTTATCCCACTATGACGCTTCTTAGAGAT 120
Qy 121 CAAAAGGGCGCTCGTGGCATCTTATCAAGTTGGCCAAAGATCTGACGCTGATGCGGCCCAT 180
Db 121 CAAAAGGGCGCTCGTGGCATCTTATCAAGTTGGCCAAAGATCTGACGCTGATGCGGCCCAT 180
Qy 181 GCGTTGGCGTTCCTCACTCGAGTTTCGGAGACACAGCTGGAGCAGTGTGCGCTTCAAC 240
Db 181 GCGTTGGCGTTCCTCACTCGAGTTTCGGAGACACAGCTGGAGCAGTGTGCGCTTCAAC 240
Qy 241 CTCCTCATGCTGGCGCTGTGTGTCAGTGGGCAATCTCTGCGAGCGCTTCCTGAGCCAG 300
Db 241 CTCCTCATGCTGGCGCTGTGTGTCAGTGGGCAATCTCTGCGAGCGCTTCCTGAGCCAG 300
Qy 301 TTCCCTTCTGGGAAGTGGTCACTCACTGTTTCACTATTCGGCTGCGCCACCATGAGTGCT 360
Db 301 TTCCCTTCTGGGAAGTGGTCACTCACTGTTTCACTATTCGGCTGCGCCACCATGAGTGCT 360
Qy 361 TTGTCGGTGTGATCTCAGTGGATGCTGTCTGGGGAGAGTCAACTTGGCGCAGTTGGTG 420
Db 361 TTGTCGGTGTGATCTCAGTGGATGCTGTCTGGGGAGAGTCAACTTGGCGCAGTTGGTG 420
Qy 421 GTGATGCTGCTGTGGAGTGACAGCTTAGCAACCTGAGCACTGAGTGTGATAGTAAATTC 480
Db 421 GTGATGCTGCTGTGGAGTGACAGCTTAGCAACCTGAGTGTGATAGTAAATTC 480
Qy 481 TTCAACACAGACTACCAATGAACATGATGCACTCTAGCTTTCGAGCTTATTTGGG 540
Db 481 TTCAACACAGACTACCAATGAACATGATGCACTCTAGCTTTCGAGCTTATTTGGG 540
Qy 541 CTGTCGTGGCTGTGCTGCTCAAGCCTCTACCCGAGGGAACGAGGATAAAGATCAG 600
Db 541 CTGTCGTGGCTGTGCTGCTCAAGCCTCTACCCGAGGGAACGAGGATAAAGATCAG 600
Qy 601 ACAGCAACGATACCCAGTTTGTCTGCAAGTGTGCGGCGCCCTCTTCTTGTGGATCTTCTGG 660
Db 601 ACAGCAACGATACCCAGTTTGTCTGCAAGTGTGCGGCGCCCTCTTCTTGTGGATCTTCTGG 660
Qy 661 CCAAGTTTCACTCTGCTCTGTGAGAGTCCAAATCGAAAGGAAGAAATGCCGTGTTCAAC 720
Db 661 CCAAGTTTCACTCTGCTCTGTGAGAGTCCAAATCGAAAGGAAGAAATGCCGTGTTCAAC 720
Qy 721 ACCTACTATGCTGTAGCAGTCAAGCGTGTGACAGCATCTCAGGGTCACTCTTGGCTCAC 780
Db 721 ACCTACTATGCTGTAGCAGTCAAGCGTGTGACAGCATCTCAGGGTCACTCTTGGCTCAC 780
Qy 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCAAGTGTGCGGCTTTCGAGGAGCGGTGGCT 840
Db 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCAAGTGTGCGGCTTTCGAGGAGCGGTGGCT 840
Qy 841 GTGGGTACCTGTGTCACTGATCCCTTCTCGTGGCTTGGCCATGGTGGGTCTTGTG 900
```

Db 841 GTGGGTACCTCGTGTCACTGATCCCTTCTCCGTGGCTTGCCATGGTGTGGCTTTGTG 900  
Qy 901 GCTGGGCTGATCTCCGTGGGGAGCCAAAGTACCTGCGGGGTGTGTAAACCGAGTGTG 960  
Db 901 GCTGGGCTGATCTCCGTGGGGAGCCAAAGTACCTGCGGGGTGTGTAAACCGAGTGTG 960  
Qy 961 GGGATTCCCAACAGCTCCATCATGGGTACAACTTTCAGCTTGTGGTGTCTCTTGGAGAG 1020  
Db 961 GGGATTCCCAACAGCTCCATCATGGGTACAACTTTCAGCTTGTGGTGTCTCTTGGAGAG 1020  
Qy 1021 ATCATCTACATGTGTGCTGTGCTTGATACCTCGGAGCCGCAATGGCATGATTTGGC 1080  
Db 1021 ATCATCTACATGTGTGCTGTGCTTGATACCTCGGAGCCGCAATGGCATGATTTGGC 1080  
Qy 1081 TTCCAGTGTCTTCACAGATGGGAACTCAGCTTGGCCATCGTGATAGCTCTCACTCT 1140  
Db 1081 TTCCAGTGTCTTCACAGATGGGAACTCAGCTTGGCCATCGTGATAGCTCTCACTCT 1140  
Qy 1141 GGTCTCTGACAGTGTTCCTCTAAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200  
Db 1141 GGTCTCTGACAGTGTTCCTCTAAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200  
Qy 1201 TATTTTGTATGACCAAGTTTCTGGAAGTTCCTCATTTGGCTGTGGATT 1251  
Db 1201 TATTTTGTATGACCAAGTTTCTGGAAGTTCCTCATTTGGCTGTGGATT 1251

## RESULT 12

AY449382 1251 bp mRNA linear PRI 03-DEC-2003  
LOCUS Homo sapiens rhesus blood group D antigen DLO mRNA, partial cds.  
DEFINITION  
ACCESSION AY449382  
VERSION AY449382.1 GI:38565512  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1251)  
AUTHORS Noizat-Pirenne, P. and Ansart-Pirenne, H.  
TITLE Direct Submission  
JOURNAL Submitted (15-OCT-2003) Blood Bank, EFS, 151 Avenue du Marechal de  
Lattre de Tassigny, Creteil 94010, France

FEATURES  
source

1. .1251  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1. .>1251  
/note="weak D DLO"  
/codon\_start=1  
/product="rhesus blood group D antigen DLO"  
/protein\_id="AA824079.1"  
/db\_xref="GI:38565513"  
/translations="MSKYPKSVRRCLPLWALFLEAAILLLFFTHYDASLEDKQL  
VASYQVQDLTVMAATLGLFLTSFRHSVSNVAFNLFLMALGVOWAILLDPLSFP  
SGKVIITLFSIRLATMSALSVLSVDVLKVNLAQVLMVLVEVTALGNLRWVIGNI  
FNTDYEMNMWHIYVPAAYFGLSVAVCLPKPLPEGTEDKDOTATIPSLSAMLGALFLWI  
FWPSFNALRSLPIERKNAVNTYAYAVSVVTAISGSLAHPQGIKSYVHSVAVIA  
GGVAVCTLHLIPSPMLAMVLGLVAGLISVGGAKYLPGCCNRVLGPHSSIMGINFSL  
LGILLGIIYIVLLVDTVGANGMIGFQVLLSISGLSLAIVIALTSLGLTGLLLNLKI  
WKAPHEAKYFDQDQVFWKFPPLHVGPF"

## ORIGIN

Query Match 99.5%; Score 1247.8; DB 9; Length 1251;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGAGCTCTAAGTACCGCGGTGTGTCGGCGTGTGCTGCTTGGGCCCTTAACACTG 60  
Db 1 ATGAGCTCTAAGTACCGCGGTGTGTCGGCGTGTGCTGCTTGGGCCCTTAACACTG 60  
Qy 61 GAAGCAGCTCTCATCTCTCTCTCTATTTTACCACATATGACGCTTCCTTAGAGAT 120

Db 61 GAAGCAGCTCTCATCTCTCTCTCTATTTTACCACATATGACGCTTCCTTAGAGAT 120  
Qy 121 CAAAGGGGCTCGTGGCATCTCTCAAGTTGGCCAAAGATCTGACCGTGTATGGCGCCATT 180  
Db 121 CAAAGGGGCTCGTGGCATCTCTCAAGTTGGCCAAAGATCTGACCGTGTATGGCGCCATT 180  
Qy 181 GGCTTGGGCTTCTCAGCTCAGCTTCGGGAGACACAGCTGGAGCACTGTGGCCCTTCAAC 240  
Db 181 GGCTTGGGCTTCTCAGCTCAGCTTCGGGAGACACAGCTGGAGCACTGTGGCCCTTCAAC 240  
Qy 241 CTCTTCATGCTGGCGCTTGTGTGCACTGGGCAATCTCTGGGAGCGCTTCTGAGCCAG 300  
Db 241 CTCTTCATGCTGGCGCTTGTGTGCACTGGGCAATCTCTGGGAGCGCTTCTGAGCCAG 300  
Qy 301 TTCCCTTCTGGGAAGTGGTCAACACTGTTCAGTATTCGGCTGGCCACCATGAGTGTCT 360  
Db 301 TTCCCTTCTGGGAAGTGGTCAACACTGTTCAGTATTCGGCTGGCCACCATGAGTGTCT 360  
Qy 361 TTGTGGTGTGATCTCAGTGGATGTCTTGGGAGGTCACCTTGGGAGGTCAGTTGGTG 420  
Db 361 TTGTGGTGTGATCTCAGTGGATGTCTTGGGAGGTCACCTTGGGAGGTCAGTTGGTG 420  
Qy 421 GTGATGTGTGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGGTCACTAGTAATATC 480  
Db 421 GTGATGTGTGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGGTCACTAGTAATATC 480  
Qy 481 TTCAACACAGACTACCAATGAAATGATGATGATGATGATGATGATGATGATGATGATG 540  
Db 481 TTCAACACAGACTACCAATGAAATGATGATGATGATGATGATGATGATGATGATGATG 540  
Qy 541 CTGCTGTGGCTGTGGTCCCTGCAAGACCTTACCGGAGGAAACGAGGATAAGATCAG 600  
Db 541 CTGCTGTGGCTGTGGTCCCTGCAAGACCTTACCGGAGGAAACGAGGATAAGATCAG 600  
Qy 601 ACAGCAACGATACCCAGTTTGTCTGCACTGCTGGGCGCCCTCTTCTTGTGGATGTTCTGG 660  
Db 601 ACAGCAACGATACCCAGTTTGTCTGCACTGCTGGGCGCCCTCTTCTTGTGGATGTTCTGG 660  
Qy 661 CCAAGTTTCACTCTGCTGTGAGAGTCCAAATCGAAAGAGAAATGCGGTGTTCAAC 720  
Db 661 CCAAGTTTCACTCTGCTGTGAGAGTCCAAATCGAAAGAGAAATGCGGTGTTCAAC 720  
Qy 721 ACCTACTATGCTGTAGCAGTGTGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAG 780  
Db 721 ACCTACTATGCTGTAGCAGTGTGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAG 780  
Qy 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCACTGCGGTGTTGGCAGGAGCGTGGCT 840  
Db 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCACTGCGGTGTTGGCAGGAGCGTGGCT 840  
Qy 841 GTGGGTACCTCGTGTCACTGATCCCTTCTCCGTGGCTTGCATGCTGGTGTGCTGTG 900  
Db 841 GTGGGTACCTCGTGTCACTGATCCCTTCTCCGTGGCTTGCATGCTGGTGTGCTGTG 900  
Qy 901 GCTGGGCTGATCTCCGTGGGGAGCAAGTACTCTCCGGGGGTGTTGTAACCGAGTGTG 960  
Db 901 GCTGGGCTGATCTCCGTGGGGAGCAAGTACTCTCCGGGGGTGTTGTAACCGAGTGTG 960  
Qy 961 GGGATTCCCAACAGCTCCATCATGGGTACAACTTTCAGCTTGTGGTGTCTCTTGGAGAG 1020  
Db 961 GGGATTCCCAACAGCTCCATCATGGGTACAACTTTCAGCTTGTGGTGTCTCTTGGAGAG 1020  
Qy 1021 ATCATCTACATGTGTGCTGTGCTTGTATACCGTGGAGCGGCAATGGCATGATTTGGC 1080  
Db 1021 ATCATCTACATGTGTGCTGTGCTTGTATACCGTGGAGCGGCAATGGCATGATTTGGC 1080  
Qy 1081 TTCCAGGTCCTCTCAGCAATGGGAACTCAGCTTGGCCATCGTGTAGTAGCTCTCACTGT 1140  
Db 1081 TTCCAGGTCCTCTCAGCAATGGGAACTCAGCTTGGCCATCGTGTAGTAGCTCTCACTGT 1140  
Qy 1141 GGTCTCTGACAGTGTGCTCTTAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200

```
Db 1141 GGTCTCTGACAGGTTTGTCTCTTAATCTTAAATATGAAGACACCTCATGAGCTAAA 1200
QY 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCAATTTGGCTGTGGATT 1251
Db 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCAATTTGGCTGTGGATT 1251

RESULT 13
AY449384
LOCUS Homo sapiens rhesus blood group D antigen DLA mRNA, partial cds.
DEFINITION Homo sapiens rhesus blood group D antigen DLA mRNA, partial cds.
ACCESSION AY449384
VERSION AY449384.1 GI:38565514
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1251)
AUTHORS Noizat-Pirenne, F. and Ansart-Pirenne, H.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2003) Blood Bank, BFS, 151 Avenue du Marechal de
Lattre de Tassigny, Creteil 94010, France
FEATURES
source
1..1251
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..->1251
/note="weak D DLA"
/codon_start=1
/product="rhesus blood group D antigen DLA"
/protein_id="AAR24080.1"
/db_xref="GI:38565515"
/translation="MSSKYPKRRCLPLWALTLEAALLFYFFHYDASLEOKGL
VASVQVODLTVAIAIGLFTSSPRHSWSVAENFMAALGVOWAILLDGFLSOPP
SGKVITLFSRLTMSALSVLISVDVLGKVNLAQLVVMVLVEVTALGNRMVLSNI
ENTYHMMWMLHYFAAAYFGLSVACLPLPEGTEDKQDRIATPSLSAMLCALPLWI
FWPSNALLSPPIERKNVNTYVAVSVVVTATSGSLAHPOKISKTYVHSAVLA
GVAVTGCHLIPSPMLAMVLGLVAGLISVGAKYLPCCNRVLGIPHSINGYNFSL
LGLRLIYIVLLVLDITVGAGNMIGFQVLLSIGLSLAIVIALTSGLLTGLLLNLKI
WKAPHEAKYFDDQVFWKPHLAVGF"

ORIGIN
Query Match 99.5%; Score 1247.8; DB 9; Length 1251;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTCTAAGTACCGGGTCTGTCGGGGCTGCGTCCCTCTGGGCCCTAACACTG 60
Db 1 ATGAGCTCTAAGTACCGGGTCTGTCGGGGCTGCGTCCCTCTGGGCCCTAACACTG 60
QY 61 GAAGCAGCTCTCATTTCTCTCTTCTATTTTATCCCACTATGAGCTTCTTTAGAGAT 120
Db 61 GAAGCAGCTCTCATTTCTCTCTTCTATTTTATCCCACTATGAGCTTCTTTAGAGAT 120
QY 121 CAAAAGGGGCTCGTGGCATCTATCAAGTTGCCAAGATCTGACCGTGTATGCGGCCATT 180
Db 121 CAAAAGGGGCTCGTGGCATCTATCAAGTTGCCAAGATCTGACCGTGTATGCGGCCATT 180
QY 181 GGCCTTGGGCTTCTCACCTCGAGTTTTCGGGACACACAGCTGAGCAGTGTGGCTTCAAC 240
Db 181 GGCCTTGGGCTTCTCACCTCGAGTTTTCGGGACACACAGCTGAGCAGTGTGGCTTCAAC 240
QY 241 CTCCTTCATGCTGGCGTGTGTCAGTGGGCAATCTCTGTCGAGCGGTTCCTGAGCCAG 300
Db 241 CTCCTTCATGCTGGCGTGTGTCAGTGGGCAATCTCTGTCGAGCGGTTCCTGAGCCAG 300
QY 301 TTCCTTCTGGGAAGTGTGTATCATCTGTTTCAAGTATTCGGCTGCCACCATGATGCT 360
Db 301 TTCCTTCTGGGAAGTGTGTATCATCTGTTTCAAGTATTCGGCTGCCACCATGATGCT 360
QY 361 TTGTCGGTGTGATCTCAGTGGATGCTGTCCTTGGGGAAGTCAACTTGGCGAGTTGGT 420
```

```
Db 361 TTGTCGGTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGAGTTGGT 420
QY 421 GTGATGGTGGTGGAGGTGACAGCTTTAGCAACCTGAGGATGTCATCAGTAATATC 480
Db 421 GTGATGGTGGTGGAGGTGACAGCTTTAGCAACCTGAGGATGTCATCAGTAATATC 480
QY 481 TTCAACACAGACTACACCATGAAACATGATGACATCTACCTGTCGACGCTATTTGGG 540
Db 481 TTCAACACAGACTACACCATGAAACATGATGACATCTACCTGTCGACGCTATTTGGG 540
QY 541 CTGCTGTGGCTGGTGCCTGCCAAAGCTCTACCCGAGGGAACGGAGGATAAAGATCAG 600
Db 541 CTGCTGTGGCTGGTGCCTGCCAAAGCTCTACCCGAGGGAACGGAGGATAAAGATCAG 600
QY 601 ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGGGCGCCCTCTTCTTGGATGTTCTGG 660
Db 601 ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGGGCGCCCTCTTCTTGGATGTTCTGG 660
QY 661 CCAAGTTTCAACTCTGCTCTGCTGAGAAATGCAATCGAAAGAAAGATGCCGTGTTCAAC 720
Db 661 CCAAGTTTCAACTCTGCTCTGCTGAGAAATGCAATCGAAAGAAAGATGCCGTGTTCAAC 720
QY 721 ACCTACTATGCTGTAGCAGTCTGAGTGTGACAGCATCTCAGGCTCATCTTGGCTCAC 780
Db 721 ACCTACTATGCTGTAGCAGTCTGAGTGTGACAGCATCTCAGGCTCATCTTGGCTCAC 780
QY 781 CCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGTGGCAGGAGGCGTGGCT 840
Db 781 CCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGTGGCAGGAGGCGTGGCT 840
QY 841 GTGGGTACCTCGTGTACCTGATCCCTTCTCGTGGCTTGCATGGTGTGCTTGTG 900
Db 841 GTGGGTACCTCGTGTACCTGATCCCTTCTCGTGGCTTGCATGGTGTGCTTGTG 900
QY 901 GCTGGGTGATCTCCGTGGGGAGCCAAAGTACCTGCGGGGTGTGTAAACGAGTGTG 960
Db 901 GCTGGGTGATCTCCGTGGGGAGCCAAAGTACCTGCGGGGTGTGTAAACGAGTGTG 960
QY 961 GGGATTTCCCACAGCTCCATCATGGGCTCAACTTTCAGTTTGTGGGTCTGCTTGGAGAG 1020
Db 961 GGGATTTCCCACAGCTCCATCATGGGCTCAACTTTCAGTTTGTGGGTCTGCTTGGAGAG 1020
QY 1021 ATCATCTACATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 ATCATCTACATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 TTCAGGTCCTCTCAGCATTTGGGAACTCAGCTTGGCCATCTGATAGCTCTCACGCTCT 1140
Db 1081 TTCAGGTCCTCTCAGCATTTGGGAACTCAGCTTGGCCATCTGATAGCTCTCACGCTCT 1140
QY 1141 GGTCTCTGACAGGTTTGTCTCTCTAAATCTTAAATATGAAAGACACCTCATGAGCTAAA 1200
Db 1141 GGTCTCTGACAGGTTTGTCTCTCTAAATCTTAAATATGAAAGACACCTCATGAGCTAAA 1200
QY 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCAATTTGGCTGTGGATT 1251
Db 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCAATTTGGCTGTGGATT 1251

RESULT 14
AY449384
LOCUS Homo sapiens rhesus blood group D antigen DRO mRNA, partial cds.
DEFINITION Homo sapiens rhesus blood group D antigen DRO mRNA, partial cds.
ACCESSION AY449384
VERSION AY449384.1 GI:38565516
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1251)
AUTHORS Noizat-Pirenne, F. and Ansart-Pirenne, H.
```





*This Page Blank (uspto)*